Reaction Mechanisms of Metalloenzymes and Synthetic Model Complexes Activating Dioxygen
Reaction Mechanisms of Metalloenzymes and Synthetic Model Complexes Activating Dioxygen

A Computational study

Valentin Georgiev
I dedicate this thesis to my wife Polina and our kids Monica and Filip, who are the best thing in my life.
Quantum chemistry has nowadays become a powerful and efficient tool that can be successfully used for studies of biosystems. It is therefore possible to model the enzyme active-site and the reactions undergoing into it, as well as obtaining quite accurate energetic profiles. Important conclusions can be drawn from such profiles about the plausibility of different putative mechanisms.

Density Functional Theory is used in the present thesis for investigation of the catalytic mechanism of dioxygenase metallo-enzymes and synthetic model complexes. Three enzymes were studied – Homoprotocatechuate 2,3-dioxygenase isolated from *Brevibacterium fuscum* (*Bf* 2,3-HPCD), Manganese-Dependent Homoprotocatechuate 2,3-Dioxygenase (*MndD*) and Homogentisate Dioxygenase (*HGD*). Models consisting of 55 to 208 atoms have been built from X-ray crystal structures and used in the calculations. The computed energies were put in energy curves and were used for estimation of the feasibility of the suggested reaction mechanisms. A non-heme \([\text{L4Me4}Fe(\text{III})]^3\) complex that mimics the reactivity of intradiol dioxygenases, and a heme \([T(o-\text{Cl})PPFe]\) complex catalyzing the stepwise oxidation of cyclohexane to adipic acid, were also studied.

For the enzymes and the non-heme biomimetic complex the reaction was found to follow a mechanism that was previously suggested for extradiol and intradiol dioxygenases – ordered substrates binding and formation of peroxo species, which further undergoes homolytic O-O bond cleavage. Different reaction steps appear to be rate limiting in the particular cases: proton transfer from the substrate to the peroxide in *Bf* 2,3-HPCD, the formation of the peroxo bridge in *HGD* and the biomimetic complex, and notably, spin transition in *MndD*.

The catalytic oxidation of cyclohexane to adipic acid in the presence of molecular oxygen as oxidant was studied, a reaction of great importance for the chemical industry. Reaction mechanism is suggested, involving several consecutive oxidative steps. The highest calculated enthalpy of activation is 17.8 kcal/mol for the second oxidative step.
List of Papers

This thesis is based on the following papers:


IV. Valentin Georgiev, Holger Noack, Margareta R. A. Blomberg and Per E. M. Siegbahn, A DFT Study on the Catalytic Reactivity of a Functional Model Complex for Intradiol-Cleaving Dioxygenases. *In manuscript*

V. Holger Noack, Valentin Georgiev, Johannes Adam Johannson, Margareta R. A. Blomberg and Per E. M. Siegbahn, The Conversion of Cyclohexane to Adipic Acid catalyzed by an Iron-Porphorin Complex. A theoretical study. *In manuscript*

Reprints were made with permission from the publishers.
Comments on the Contribution to the Papers

I have performed the all the calculations and prepared the manuscripts for Paper II and Paper III. For Paper I I was involved in the discussion, and I have performed additional calculations for testing the concerted Criegee rearrangement. For Paper IV I have performed all the calculations on the suggested intradiol mechanism and prepared the manuscript except for the Introduction. The calculations on the alternative extradiol path were done by Holger Noack, who prepared the Introduction part. For Paper V I have done all the calculations for the first and the third major reaction steps of the suggested mechanism, and prepared the manuscript. Holger Noack performed calculations on alternative reaction paths. Johannes Johansson calculated the second main reaction step and contributed to discussion about entropy effects in the manuscript.
Contents

Part I: Catalysis, dioxygen, and iron
1 How do the enzymes function ........................................ 15

Part II: Theory and methods
2 Quantum Chemistry ................................................... 21
3 Density Functional Theory ........................................... 27
4 Transition State Theory ............................................... 31
5 Technical details ...................................................... 33

Part III: Dioxygenases
6 Extradiol and intradiol catechol dioxygenases .................. 39
7 Fe- and Mn-dependent homoprotocatechuate 1,2-dioxygenas es . 43
   7.1 Background .......................................................... 43
   7.2 The model ............................................................ 46
   7.3 Reaction mechanism of Bf 2,3-HP CD .......................... 46
   7.4 Reaction mechanism of MndD. Spin transition .............. 49
8 Homogentisate Dioxygenase ......................................... 57
   8.1 Background .......................................................... 57
   8.2 The model ............................................................ 58
   8.3 Enzyme-Substrate Complex ...................................... 59
   8.4 Reaction mechanism of Homogentisate Dioxygenase ....... 59
9 Summary .................................................................... 63

Part IV: Biomimetic model complexes
10 Non-heme iron model complex for intradiol-cleaving dioxygenases .................. 67
   10.1 Background .......................................................... 67
   10.2 The model ............................................................ 67
   10.3 Reaction mechanism ............................................... 69
   10.4 Intra- vs. Extradiol bond cleavage ............................. 73
   10.5 Comparison with the biological system ....................... 74
11 Heme complex catalyzing adipic acid synthesis ................. 77
   11.1 Background .......................................................... 77
   11.2 Computational technique ......................................... 78
   11.3 Suggested reaction mechanism ................................... 80
      11.3.1 Oxidation of cyclohexane to 1,2-cyclohexanediol ....... 80
      11.3.2 1,2-cyclohexanediol oxidation and ring cleavage ....... 83
11.3.3 Oxidation of adipic aldehyde to adipic acid ......................... 85

Part V: Concluding Remarks

Part VI: Acknowledgements

Part VII: Populärvetenskaplig sammanfattning på svenska

Bibliography ................................................................. 97
Part I:
Catalysis, dioxygen, and iron

Chemical reactions can be fairly slow at ambient conditions, and the desired products are often accompanied by side products. This not only means low efficiency, but danger as well, since the side products can be hazardous and even toxic. Catalysts in general are substances which are able to accelerate chemical reactions without changing the chemical equilibrium and without being consumed.
1. How do the enzymes function

The biological catalysts found in Nature are known as enzymes. Almost all known enzymes are protein molecules, but some RNA molecules (ribozymes and synthetic deoxyribozymes) also hold catalytic properties. The protein macromolecules consist of long chains of amino acids bound together by peptide bonds [1]. These chains are furthermore folded specifically in the secondary structure, and thus special regions in the protein structure are formed, called active sites. The active site is the part of the enzyme that is responsible for it’s physiological role, it is the place where the catalysis occurs. Due to the specific structural, chemical and electrostatic features of the active site, normally only one type of substrate molecules can bind into it, hence the enzyme is highly selective. Having in mind the great variety of chemical reactions and therefore substrates, we can imagine the huge number of enzymes Nature has designed. The enzymes are extremely effective, enhancing the reaction rates millions of times and acting under mild conditions. Thus they catalyze biological reactions, which under normal conditions would proceed with negligible rate. Enzymes also show impressive levels of stereospecificity, regioselectivity and chemoselectivity [2].

A rough picture of an enzyme-catalyzed reaction can be described in the following way: when a substrate molecule enters the active site an Enzyme-Substrate complex (ES) is formed, from which the catalytic reaction proceeds and the product is formed. In the Enzyme-Product complex (EP) the hydrogen bonds and electrostatic interactions, previously stabilizing the ES, are destroyed and the product molecule can easily leave the active site [3]. However, the origin of the catalytic power, the means of achieving the reaction rate enhancement, and the specificity, are all widely debated phenomena without straightforward explanation.

There are number of theories trying to explain how the enzymes do their work. One of the most widely used is the *lock-and-key* theory. It was suggested in 1894 by Emil Fisher and is based on the assumption that the substrate molecule corresponds perfectly to the active site, so that once it enters the enzyme, it is oriented in a particular optimal spatial position and then “locked” into the active site with hydrogen bonds and electrostatic interactions. This is a very illustrative picture which unfortunately provides only a reasonable explanation of the specificity of the enzymes. The *induced
fit model was suggested in 1958 by Daniel Koshland [4] as a modification to the lock and key model. More up-to-date theories adopt the ideas of transition state stabilization and/or ES complex destabilization, providing an alternative pathway, reducing the reaction entropy by bringing substrates together. All of them are based on the concept of lowering the activation energy barrier $\Delta G^\ddagger$ of the reaction. Extensive and elaborate discussions on the origin of the catalytic power can be found in references [3, 5, 6].

Particularly interesting are the metal containing enzymes, also known as metalloproteins. They contain a metal ion cofactor, usually coordinated by nitrogen, oxygen or sulfur atoms from the polypeptide’s amino acids. The metal ion can be also coordinated by macrocyclic ligands incorporated in the protein, as heme, chlorophyll, and vitamin B12. The presence of the metal ion allows these enzymes to perform oxidation and reduction, reactions that are not easily achieved by the organic functional groups present in the proteins. Thus the metalloenzymes can catalyze biochemical reactions of great importance, like photosynthesis, O-O bond cleavage in cellular respiration, oxygenation, dioxygenation etc. From an industrial point of view dioxygen (and hydrogen peroxide to some extend) is the perfect oxidant - it is available in the Earth’s atmosphere, and it doesn’t cause environmental problems. Molecular oxygen adopts an "unusual" triplet electronic configuration in its ground state, which prevents it from reacting with closed shell molecules in ambient conditions. Such reactions are spin-forbidden since they involve change of the total spin state of the system. Chemists achieve activation of dioxygen at higher temperatures and pressures, quite severe conditions, which makes the utilization of these reactions expensive for the chemical industry. Nature however, utilizes transition metal ions in metalloproteins to perform the desired reactions at normal temperature and neutral pH. In biology, the proteins containing copper, heme-iron, and non-heme iron are the main catalysts involved in dioxygen activation and transportation.

In order to better understand the metalloproteins, synthetic inorganic model complexes [7] have been extensively studied, both experimentally and theoretically. The use of these biomimetic complexes gives the possibility for fine tuning of the catalytic reactivity. It is easier to create different biomimetic systems by using different ligands, solvents, counterions etc., rather than performing mutagenesis studies on real enzymes.

Understanding the enzyme structure and reaction pathways is a direct way to address the question about the nature of the enzymatic power. Such knowledge is of great importance in all areas where catalysis is involved, like drug design, pharmaceutics, chemical industry, etc. Even though there has been a great development in the science and technology in the last decades, experimental studies are usually expensive, and can be very slow as well. Regarding reaction mechanisms, there are also
limited capabilities for studying extremely short-living intermediates and transition states. Fortunately, advances in computer technology and quantum theory provide a possibility for carrying out theoretical research on enzyme catalysis. Using computer modeling one can investigate such "invisible" species as transition states, thus providing useful information which can support experimental results and help in their interpretation.
Part II: Theory and methods

The modeling of chemical reactions can be done using Quantum Mechanics (QM) implemented in the computational codes, i.e. Quantum Chemistry (QC). Other methods like Molecular Dynamics (MD) and Molecular Mechanics (MM) are based on classical mechanics, they are computationally cheap and one can use them to describe systems consisting of thousands of atoms. Thus these methods can be used for investigation of big bio-molecules like proteins as a whole – protein structure determination and prediction. Quantum methods, which allow for more accurate solving of chemical problems, are quite demanding regarding computer power and hence only systems of limited size (up to 200 atoms nowadays) can be treated.

Quantum mechanical equations give a correct description of the behavior of electrons, but they can only be solved exactly only for one-electron systems. Plenty of methods based on different approximations are available for solving many-electron problems, and choosing a relevant one for characterization of enzyme catalytic mechanisms is an important question [8, 9].

The main physical quantity calculated in a QM study is the energy of different structures of the investigated system. These energies are used for prediction of intermediates, transition states and activation barriers. The connection between energy barriers and reaction rates is given by Transition State Theory (TST). The reaction rates calculated in this way are compared with the corresponding experimental observations, and should enable to exclude or favor the suggested catalytic mechanism.
2. Quantum Chemistry

Quantum chemistry applies quantum mechanics and quantum field theory to address theoretical problems in chemistry. The quantum energy and other important properties of a given system can be obtained by solving the time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi$$  \hspace{1cm} (2.1)

where $\hat{H}$ is the Hamiltonian operator, which corresponds to the physical quantity energy, $\Psi$ is the wave function of the system, and $E$ is the corresponding total energy. The wave function is a mathematical tool which is used for describing all the properties of any physical system and it plays a central role in quantum mechanics. In general $\Psi$ is a function of coordinates and time, but in the scope of this work we are not interested in how it evolves in time and therefore we consider the time-independent Schrödinger equation. The wave function is a probability amplitude and has no physical meaning itself, but its square $|\Psi(r)|^2$ is a probability density. Multiplied by a volume element it gives the probability $|\Psi(r)|^2 dV$ that a particle can be found in the volume element $dV$ at the point $r$. In order to represent a physically observable system, the wavefunction must be continuous, single-valued and finite everywhere, and its first derivative must be continuous too. The Hamiltonian of a system consisting of $N$ electrons and $M$ nuclei has the form (in atomic units):

$$\hat{H} = -\sum_{i} \frac{1}{2}\nabla_i^2 - \sum_{A} \frac{1}{2M_A} \nabla_A^2 - \sum_{i} \sum_{A} \frac{Z_A}{r_{iA}} + \sum_{i} \sum_{j>i} \frac{1}{r_{ij}} + \sum_{A} \sum_{B>A} \frac{Z_A Z_B}{R_{AB}}$$  \hspace{1cm} (2.2)

The first two terms are the kinetic energy terms for the electrons and the nuclei, respectively. The next three terms represent the Coulomb potential arising from electron-nucleus attraction, electron-electron repulsion, and nucleus-nucleus repulsion respectively.

The wave function methods (see [8, 10] for more details and references) try to find approximate solutions directly to the Schrödinger equation. Another approach takes the energy as a functional of the electron density. Then, instead of calculating the wave function explicitly, one has to deal with the electron density, which depends only on 3 spatial variables, no matter how many electrons are present in the system. This branch of quantum chemical methods is called Density Functional Theory (DFT) [11, 12] and it will be given a short overview in the next section.
Since the wave function depends on the coordinates and the spin of the electrons, solving the Schrödinger equation means dealing with 3 spatial and 1 spin variables for each electron, 4N variables in total (where N is the number of electrons). The resulting equations can be solved exactly only for one-electron systems like the \( \text{H}_2^+ \) molecule. All the atomic and molecular systems that contain more than one electron lead to complex equations that can only be solved by applying different approximations.

One of the main approximations that significantly simplifies the solution of the Schrödinger equation for real many-electron systems is the Born-Oppenheimer approximation. It is based on the fact that the nuclei are much heavier than the electrons and their velocities are much lower. Therefore the Hamiltonian can be divided into an electronic part and a nuclear part. The electronic Hamiltonian (eq. 2.3), where the nuclear positions enter as parameters, defines the electronic Schrödinger equation, which is satisfied by the electronic wave function \( \Psi_{el} \) and the electronic energy \( E_{el} \), correspondingly.

\[
\hat{H}_{el} = -\sum_i^N \frac{1}{2} \nabla_i^2 - \sum_i^N \sum_A^M \frac{Z_A}{r_{iA}} + \sum_i^N \sum_{j>i}^N \frac{1}{r_{ij}} \tag{2.3}
\]

The total energy is then a sum of the electronic energy and a term expressing a constant nuclear repulsion:

\[
E_{tot} = E_{el} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \tag{2.4}
\]

Once the electronic problem is solved for different nuclear positions one can construct a nuclear Hamiltonian for the motion of the nuclei in the average field of the electrons. In that way the electronic energy appears as a potential on which the nuclei move – Potential Energy Surface (PES). The solutions of the nuclear Schrödinger equation correspond to translational, vibrational, and rotational states of the molecule.

Another approximation is to construct the \( n \)-particle wave function from \( n \) one-particle wave functions. There are certain conditions, which the wave function must satisfy in order to properly describe some peculiar properties of the electrons. Electrons are indistinguishable particles, which means that the wave function must allow any of them to occupy any electronic state. A spin variable must be included in the one-electron functions in order to describe the intrinsic angular momentum of electrons. As being fermions (particles with half integer spin) the electrons must be described by an antisymmetric wave function – the well known Pauli exclusion principle. The spin of the electron is taken into account by representing the one-electron wave function as a product of a spatial orbital and a spin function, resulting in the so-called spin-orbital:
\[ \chi(r, s) = \psi(r) \sigma(s) \quad (2.5) \]

The indistinguishability and Pauli principles are satisfied by a wave function constructed in the form of a Slater determinant (SD):

\[ \Psi(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{n!}} \left| \begin{array}{cccc} \chi_1(x_1) & \chi_2(x_1) & \cdots & \chi_n(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \cdots & \chi_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_n) & \chi_2(x_n) & \cdots & \chi_n(x_n) \end{array} \right| \quad (2.6) \]

where \( \chi_i(x) \) is the \( i \)th spin-orbital occupied by the \( i \)th electron. If two electrons occupy the same spin-orbital two rows of the determinant will become equal, i.e. the determinant will vanish (Pauli exclusion principle). An interchange of two rows, which corresponds to interchange of the coordinates of two electrons, changes the sign of the determinantal wave function, thus satisfying the requirement for antisymmetry.

The basic method for solving the Schrödinger equation in quantum chemistry is the Hartree-Fock (HF) method. This is the groundwork of the ab initio methods, those that solve the many-body time-independent Schrödinger equation from fundamental physical principles without using empirical parameters. The Hartree-Fock method uses the variational principle to optimize the Slater determinant as a trial wave function. The variational principle states that for a time-independent Hamiltonian operator, any trial wavefunction will have an energy that is higher than or equal to the true ground state energy corresponding to the given Hamiltonian.

In a procedure known as the Self-Consistent Field (SCF) a single Slater determinant is optimized with respect to the energy. The optimization of the SD consists of varying the spin-orbitals with the constraint that they should remain orthonormal. The spin-orbitals are solutions of the Hartree-Fock equations:

\[ \hat{f}_i \chi_i = \varepsilon_i \chi_i \quad i = 1, 2, \ldots, n \quad (2.7) \]

where \( \hat{f}_i \) is an effective one-electron operator known as the Fock operator. Apart from the already mentioned terms for the electronic kinetic energy and electron-nuclei interactions, the third term entering the Fock operator (eq. 2.8) expression is the Hartree-Fock potential \( V_i^{HF} \).

\[ \hat{f}_i = -\frac{1}{2} \nabla_i^2 - \sum_i \sum_A Z_A \frac{1}{r_{iA}} + V_i^{HF} \quad (2.8) \]

The Fock potential is the average repulsion experienced by the \( i \)th electron due to the field created by all the remaining electrons in the system. \( V_i^{HF} \) can be written in the following form:
\[ V_i^{HF}(x_1) = \sum_{j}^{N} (\hat{J}_j(x_1) - \hat{K}_j(x_1)) \] (2.9)

where \( \hat{J} \) is the **Coulomb operator**, and \( \hat{K} \) is the **Exchange operator**. The first one is simply the energy of the Coulomb interaction between an electron in orbital \( i \) with an electron in orbital \( j \). This integral is always positive, i.e. destabilizing – which is what one expects from a Coulomb repulsion between electrons.

\[ \hat{J}_j(1)\phi_i(1) = \left[ \int \phi_j^*(2) \frac{1}{r_{12}} \phi_j(2) dx_2 \right] \phi_i(1) \] (2.10)

The effect of the Exchange integral is to exchange electrons in different spin orbitals, and it vanishes for electrons with opposite spin since the spin functions are orthogonal.

\[ \hat{K}_j(1)\phi_i(1) = \left[ \int \phi_j^*(2) \frac{1}{r_{12}} \phi_i(2) dx_2 \right] \phi_j(1) \] (2.11)

There is no classical interpretation for the Exchange integral. It arises from the antisymmetric nature of the wave function and represents purely quantum effect.

The spin-orbitals entering the trial wave function are expanded in a finite set of known basis functions:

\[ \psi_i(r) = \sum_{\mu=1}^{K} c_{i\mu} \phi_\mu(r) \] (2.12)

The resulting wave function is a linear combination of the initial one-electron basis function, where the expansion coefficients are optimized by the SCF procedure. Using a larger number of basis functions will generally lead to more accurate solution, but the computational costs scale as the forth power of the number of basis functions in HF.

A major problem of the Hartree-Fock method is that it lacks a certain part of the electronic energy, referred to as **correlation energy**. The reason for this is that there are two-electron terms in the Hamiltonian which are replaced by an average repulsive potential for each electron. In other words each electron feels an average field due to the repulsion from all other electrons, instead of the explicit electron-electron interaction. Although the correlation energy represents a quite small percentage of the total energy, it is rather essential for solving chemical problems and the evaluation of relative energies.

Different more advanced wave function methods have been developed in order to take into account the correlation energy explicitly. Some of them, like
Møller-Plesset perturbation theory (MP\(n\), when \(n\) is the order of the correction), Configuration Interaction (CI) and Coupled Cluster (CC) theory, achieve this by including more determinants in the wave function. The wave function thus becomes more flexible allowing the electrons to correlate their movements in different orbitals. Multi-configurational methods like MCSCF and CASSCF also add new determinants to the wave function but optimize their orbitals together with the coefficients in front of them. CASPT2 is a multireference method using additionally perturbation theory. The accuracy of the calculations with the methods listed above is improved significantly compared to Hartree-Fock, but the computational costs are also extremely increased. In MP2 for example the computational cost scales as the fifth power of the number of basis function, and in CISD and CCSD it scales as sixth power.

The correlated wave function methods can be very accurate, and in principle can converge to the "exact" solution, but their application is limited to 10-20 atoms systems due to the enormous demands for computer power.
3. Density Functional Theory

A widely used method in quantum chemistry is the density functional theory. The foundation of this theory was laid by the works of Thomas, Fermi and Dirac on uniform electron gas (1920s). They realized that the energy of the system can be expressed as a functional of the electronic density, without dealing explicitly with the wave function. This idea was further developed and brought to success by Hohenberg and Kohn in 1964 [13, 14]. They provided a new approach for solving the Schrödinger equation for many-electron systems by proving two theorems: (i) there is a one-to-one correspondence between the electronic density $\rho(r)$ and the non-degenerate ground-state energy of the system, and (ii) the true ground-state electron density can minimize the energy functional according to the variational principle. The second theorem means that it is possible to use an energy minimization procedure, similar to the one used in wave function methods. The energy functional can be written as a sum of terms from the kinetic energy $T[\rho]$, electron-electron interactions $V_{ee}[\rho]$, and nuclear potential $V_{nn}$:

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{nn}$$

(3.1)

An important contribution to practical density functional theory was given by Kohn and Sham when they introduced the electron density expressed as a linear combination of basis functions, similar to Hartree-Fock orbitals:

$$\rho(r) = \sum_i N |\psi_i(r)|^2$$

(3.2)

The computational implementations of DFT are in many aspects similar to the procedures developed for wave function methods. In its use of a single determinant and the variational principle, the Kohn-Sham approach resembles the Hartree-Fock method discussed above. Using the Slater determinant constructed from the Kohn-Sham orbitals the kinetic energy $T$ can be expressed in the same way as in the Hartree-Fock method:

$$T_{SD}[\rho] = -\frac{1}{2} \sum_i \langle \psi_i | \nabla_i^2 | \psi_i \rangle$$

(3.3)

A major principal difference is that the Kohn-Sham equation would give the exact energies and electron density, if the correct functional was used. Apply-
ing the Kohn-Sham approach the energy functional reduces to form in which all the terms but one can be analytically expressed:

\[ E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{XC}[\rho] \]  

(3.4)

where \(T_s[\rho]\) is the kinetic energy calculated from the Kohn-Sham orbitals, \(E_{ne}[\rho]\) is the nuclei-electron interaction, \(J[\rho]\) is the classical Coulomb integral, and \(E_{XC}[\rho]\) is the exchange-correlation term. This exchange-correlation functional is not exactly known and there is no systematic way to derive it. This is actually the challenging part of DFT, since the intrinsic accuracy of the method depends solely on the quality of the functional. In spite of this, new functionals with surprising accuracy were constructed over the years and DFT became a useful method for chemical applications.

The simplest functionals based only on the electron density are known as Local Density Approximation (LDA) and Local Spin Density Approximation (LSDA) for open-shell systems. The exchange-correlation energy is written as

\[ E_{LDA}^{xc}[\rho] = \int \rho(r)\varepsilon_{xc}[\rho]dr \]  

(3.5)

where \(\rho(r)\) is the local density, which varies only slowly with \(r\), and \(\varepsilon_{xc}\) is the exchange-correlation energy per electron. \(E_{LDA}^{xc}[\rho]\) can be decomposed into exchange and correlation terms, leading to separate expressions for \(\varepsilon_x\) and \(\varepsilon_c\). The exchange term was formulated analytically for a uniform electron gas, but the correlation term was only estimated [15, 16]. LSDA methods suffer from self-interaction error and overestimate bond dissociation energies. They give results with similar accuracy as HF. Examples for LSDA functionals are Vosko-Wilk-Nusair (VWN), Perdew-Zunker (PZ81), Cole-Perdew (CP), Perdew-Wang (PW92).

The next step is to include not only the density, but also its gradient:

\[ g = \frac{|\nabla \rho|}{\rho^{4/3}} \]  

(3.6)

in the so-called gradient corrected or Generalized Gradient Approximation (GGA). Among the most prominent GGA functionals are the Perdew’s P86 [17], the LYP functional by Lee, Yang and Parr [18], and BLYP developed by Becke (exchange part) and Lee, Yang and Parr (correlation part). The GGA approximation describes chemical bonding markedly better than LSDA, with accuracy similar to MP2 [19].

The most modern and successful are the hybrid functionals, which combine functionals from LSDA, corrections from GGA, a fraction of Hartree-Fock exchange \(E_{x}^{HF}\), and the introduction of a few empirical parameters. The most
prominent functional of this generation is B3LYP, which has also been used in the present thesis. The B3LYP functional can be written as:

\[ F_{xc}^{B3LYP} = (1 - A)E_{Slater}^x + AE_{HF}^x + BE_{Becke}^x + (1 - C)E_{VWN}^c + CE_{LYP}^c \]  \hspace{1cm} (3.7)

In this functional, as many others, contributions to the exchange-correlation functional are separated into exchange and correlation parts. \( E_{Slater}^x \) is the Dirac-Slater exchange, \( E_{HF}^x \) is the Hartree-Fock exchange term, \( E_{Becke}^x \) is the gradient part of the exchange functional of Becke [20, 21, 22, 23], \( E_{VWN}^c \) and \( E_{LYP}^c \) are the correlation functionals of Vosko, Wilk, and Nusair [16] and Lee, Yang, and Parr [18] respectively. The parameters A, B and C are determined empirically by Becke [23], and have values of A=0.20, B=0.72, and C=0.81. Usually the functionals are constructed to fit systems for which there exists enough experimental data or an exact solution. Density Functional Theory can thus be called a semi-empirical method, but in a way it forms a class by itself, because once constructed, the functionals can be applied to any system. The accuracy of a theoretical method is normally tested via comparison of the computational results with experimental data. Standard G2, G3 benchmark [24] tests and their variants are among the most accurate. The G2 set includes a number of molecules for which very accurate experimental data is available. In 1998 the G3 benchmark set [25] was also introduced and has been available for testing B3LYP. The accuracy of the B3LYP calculations depends on the choice of basis set. The smallest basis set which is recommended for calculations is of double-zeta quality. The tests show [26] that this rather small basis set works quite well for geometry optimizations (Table. 3.1). The average error in the atomization energies is 5.18 kcal/mol.

<table>
<thead>
<tr>
<th>Bond Lengths [Å]</th>
<th>error</th>
<th>6-31G*</th>
<th>6-311+G(3df,2p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>0.013</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>maximum</td>
<td>0.055</td>
<td>0.039</td>
<td></td>
</tr>
<tr>
<td>Bond Angles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.62°</td>
<td>0.61°</td>
<td></td>
</tr>
<tr>
<td>maximum</td>
<td>1.69°</td>
<td>1.85°</td>
<td></td>
</tr>
<tr>
<td>Dihedral Angles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average</td>
<td>0.35°</td>
<td>3.66°</td>
<td></td>
</tr>
<tr>
<td>maximum</td>
<td>0.63°</td>
<td>6.61°</td>
<td></td>
</tr>
<tr>
<td>Atomization Energy [kcal/mol]</td>
<td>average</td>
<td>5.18</td>
<td>2.20</td>
</tr>
<tr>
<td>maximum</td>
<td>31.50</td>
<td>8.40</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Mean absolute error for DFT(B3LYP) method applied to the G2 benchmark set [24, 26, 27].
Using a large basis set, as 6-311+G(3df,2p) [24, 27] the geometry accuracy is improved only slightly, but the average error for the atomization energies is much lower, 2.20 kcal/mol (Table. 3.1).

Based on the values presented in Table. 3.1 and more recent experience [28, 29, 30] it can be concluded that results obtained from B3LYP calculations have enough accuracy for studying the energetics of enzymatic reactions. The systems of interest in the present thesis however contain transition metals, and the benchmark results discussed above do not include data for such systems. Investigations of M-R (M is a first row transition metal and R is H, CH₂, CH₃ [31] or OH [32]) bond strengths in small cationic systems show an average absolute error of 4.9 kcal/mol and a maximum error of 9.0 kcal/mol for B3LYP. This performance can be considered as quite good having in mind the properties of the transition metals. The presence of degenerate 3d, 4s and 4p levels makes them hard to investigate computationally. A modified version of the B3LYP functional, which includes 15% Hartree-Fock exchange (20% in the original functional), is denoted as B3LYP*. In cases of transition metal complexes such a decrease of the amount of Hartree-Fock exchange can give a better description of the system [33]. Therefore the B3LYP* functional has been used extensively in the later work presented in this thesis. In general, states with larger exchange contribution to the total energy (high-spin states), are destabilized by B3LYP* due to the reduced amount of the exact exchange. Finally, B3LYP was found useful for studying enzyme reaction mechanisms due to its adequate accuracy and the speed it provides. The problems addressed in the present thesis consist of comparing alternative hypothetical mechanisms of certain enzyme reactions. B3LYP calculations give a lot of data with satisfactory accuracy, therefore the different hypothetical mechanisms can be evaluated and those with the highest activation barriers can be rejected.
4. Transition State Theory

The subject of investigation presented in this thesis is enzymatic chemical reactions in which the reactants are transformed into products. The mechanisms of these transformations usually pass through several steps involving different chemically stable and unstable structures (intermediates, transition states, etc.). Each step is characterized by an unstable structure, connecting the reactant and the product, and known as the Transition State (TS). This TS structure is unstable by means of the higher energy it has \( G^\ddagger \), therefore it appears to determine an energy barrier which the system should overcome in order to complete the reaction step. The activation energy barrier \( \Delta G^\ddagger \) is defined as the difference in the free energies of the reactant and the transition state. The main approach in the theoretical study of enzyme catalyzed reaction mechanisms is to calculate the Potential Energy Surface along the reaction coordinate, which involves locating and characterizing the minima and transition states and their relative energies.

The Transition State Theory (TST) \([34, 35, 36]\) makes the connection between the calculated data (relative free energies) and the reaction rates which come out from experimental measurements. TST postulates that all stable and unstable states along the reaction coordinate obey the Boltzmann energy distribution law. Based on this equilibrium energy distribution the following expression gives the connection between the rate constant \( k \), the temperature and the reaction barrier:

\[
k = \frac{k_B T}{h} \exp \left( -\frac{\Delta G^\ddagger}{RT} \right), \tag{4.1}
\]

where \( k_B \) is the Boltzmann constant; \( h \) is the Planck constant; \( T \) is the absolute temperature; \( R \) is the universal gas constant; and \( \Delta G^\ddagger \) corresponds to the free energy of activation. The exponential Eyring equation presented above gives a relation between the experiment and the theory. Experiments measure rate constants and calculations produce the energy barriers.

According to the Eyring equation, rate constant of 1 s\(^{-1}\) at room temperature corresponds to a barrier of 18 kcal/mol. One order of magnitude of the rate constant value corresponds to 1.4 kcal/mol energy barrier. Considering the expected error of B3LYP of 3-5 kcal/mol the picture seems not quite optimistic. The difference in the activation barriers of the hypothetical mechanisms however, appears to be large enough, often of the order of 10 to 30 kcal/mol. Theoretical research thus gives a unique possibility to address such
questions, since short-living species like the transition states are accessible for studies.
5. Technical details

The first step in solving the catalytic mechanism is constructing a good model of the system under investigation. The systems studied experimentally are usually too large for a quantum-chemical treatment. Therefore small models of about one hundred atoms need to be constructed. The goal is to build a small model, that is good enough to represent a system consisting of thousands of atoms.

X-ray crystal structures of huge number of enzymes are available at different databases, one of the most popular being the Protein Data Bank (PDB). These crystal structures can serve as a good starting point for the modeling process. The experience of many researchers in this area shows that in many cases including only the active site residues, the substrate molecule, and the cofactor molecules (in case there are such) into the computational model is enough for an accurate description of the chemistry occurring. The active site on the other hand consists of a number of amino acid residues and it is often impossible to include all of them in the model. In fact it is common that only few of them are directly involved in the catalysis. Another aspect of the modeling is that these few important amino acid residues can be modelled themselves and thus reduced to smaller molecules. All the amino acids have analogous structure – a central, $\alpha$-carbon atom (or $C^\alpha$), amino and carboxyl groups attached to it, a hydrogen atom and a side chain ($R$) attached also to the $\alpha$-carbon. It is enough to keep only the side chain in the model and remove the rest of the residue. Usually even the side chains are modelled using suitable organic molecules, like imidazole for Histidine, acetate for Glutamine, phenol for Tyrosine, etc. Tyrosine is even modeled as water in the present calculations for representing its hydrogen-bonding with the substrate. This technique is known to give quite adequate results for the purposes of studies of this type [37].

Although such models include explicitly the part of the enzyme where the chemical reaction takes place, the effects of the rest of the protein are certainly missing. Two techniques have been utilized in the investigations presented here in order to account for the effects of the protein environment: (i) geometry constraints in the optimization, and (ii) calculations of the solvation effects of the protein environment. The latter will be discussed in the last subsection. During the geometry optimization constraints were applied on the terminal atoms in order to fix them to their crystallographically observed positions. This approach turned out to be quite useful especially for the second-shell
residues. As they are not directly coordinated to the active site metal center they tend to move quite a lot around their original positions, which should not be the real situation in the enzyme. Introduced by Pelmenschikov [38], this technique is now widely used in our group in such kind of studies.

For the biomimetic synthetic complexes the situation is slightly different, and not at all less complicated. When the experimentally used ligand is too big some of its substituents can be removed from the model. Usually these are bulky hydrocarbon groups (ethyl, propyl, tert-butyl, etc.). Often the reacting chemical species can be completely included in the model, which means no truncation of atoms/groups, and no need for geometry constraints. This seems to be an advantage over the enzymatic modelling, but in fact the theoretical study of catalytic reactions in solution is quite hard. For example, it is not straightforward to take into account explicit interactions with the solvent molecules, couterions in the solution, base/acid molecules that are usually present in the reaction environment. As a result, the biomimetic complexes models are often charged, which complicates the modelling of the solvent effects. Processes where protons enter or leave the reaction site are also hard to model. In the enzymes all these characteristic chemical features are "brought together" in the active site by the protein structure.

Once the model is constructed, different reaction mechanisms can be probed by finding all the intermediates and transition states along the putative reaction coordinate, and plotting the energy diagram.

For all models described in this thesis, the geometry optimizations were performed using the B3LYP method as implemented in the Jaguar [39] program package. The optimizations were performed in gas phase with a double-$\zeta$ quality basis set (lacvp in Jaguar), which includes an Effective Core Potential (ECP) on iron [40]. According to the benchmark data (Table 3.1) this double zeta basis set is accurate enough in terms of geometric structures. The complicated and time consuming part of any theoretical study of chemical reactions is locating the TS. By systematically freezing one or more internal coordinates and optimizing the other degrees of freedom, the reaction coordinate can be scanned. The energy maximum in the scanned surface provides a good guess for the TS, and this structure is put in the TS optimization job. The TS is defined as a structure which has a minimal energy with respect to all coordinates except one – the reaction coordinate. Along the reaction coordinate the energy takes maximum at that point and thus the second derivative in the Hessian is negative, which corresponds to an imaginary vibrational frequency. In this sense the imaginary frequency represents movement along the reaction coordinate. Also, calculation of the molecular Hessian for the reactant, the TS, the intermediates, and the products, are used to estimate the Zero-Point vibration Energy (ZPE), and the entropy contributions to the total energy. Entropy effects cannot be described accurately when some coordinates in the model
are fixed, therefore they are not included in the results for the enzymatic systems. These effects should play a minor role for the chemistry in the course of the reaction, but will be important when the dioxygen molecule binds at the active site. In the present thesis, frequency calculations and transition state optimizations have been performed using Gaussian03 [41], at the B3LYP/lacvp level of theory.

As it was stated above the B3LYP geometries are not very sensitive to the basis set quality [29]. Relative energies on the other hand are particularly much more sensitive to the basis set, therefore it is a usual practice to apply larger basis sets for obtaining accurate energies. In the present studies this is done by performing single point calculations on the lacvp-optimized geometries using triple-\(\zeta\) basis set – \(cc-pVTZ(-f)\) (without f-functions) in Jaguar. This correlation-consistent polarized basis set, which is intrinsically polarized but does not include ECP, was used for all atoms except iron. For iron the triple zeta quality basis set \(lacv^3p\) was explicitly used.

The DFT method is limited to a single determinant. Open-shell low-spin states are linear combinations of different spin determinants and cannot be described accurately by a single determinant. The calculated low-spin states are therefore spin contaminated by higher spin states. The broken symmetry approach based on the Heisenberg Hamiltonian and introduced by Noodleman [42, 43] was used to correct the energy of the low-spin states.

Geometry optimizations and the large basis sets single point calculations are performed in gas phase. This model is apparently missing the effects of the protein environment where the enzymatic reactions occur. An important part in such an investigation is to consider the effect of the environment. A common way to compute the influence of the surrounding environment is by performing a single point calculation using the Polarizable Continuum Model (PCM), as implemented in Jaguar [44, 45].

A dielectric constant (\(\varepsilon\)) equal to 4.0 is used to set up the surrounding environment in the studies of enzymatic reactions. This value has been empirically determined, and it reproduces an environment resulting from a dielectric constant of about 3 for the protein itself and about 80 for the water medium surrounding the protein [46]. When modeling the biomimetic complexes, the dielectric constant of the solvent used in the experiment was applied in the calculations. For the intradiol cleaving non-heme iron complex was used a value of \(\varepsilon=37.5\), corresponding to acetonitrile. For the adipic acid synthesis reaction \(\varepsilon=2.0\) was used, which is the dielectric constant for cyclohexane.

Usually, the dielectric effects are not expected to change the relative energies dramatically, except in cases of reactions where charge separation occurs [47, 48]. Large effects are considered to be a sign of a problem with the chosen model. The dielectric continuum model cannot account for short-range solute-solvent interactions such as hydrogen bonds [49, 50, 51, 52]. Therefore
one should have in mind that it gives a kind of qualitative results, and it is always better to include hydrogen bonds explicitly in the model, if possible.
Part III:
Dioxygenases

The last part of the catabolism of aromatic compounds in the environment involves cleavage of the aromatic ring and insertion of both dioxygen atoms into the product, a reaction known as dioxygenation. In Nature such spin-forbidden reactions with participation of molecular oxygen, are catalyzed by enzymes containing transition metal ions as cofactors. Transition metals have low-lying electronic states with unpaired electrons, and very often they have high-spin electronic ground states. The large spin-orbit coupling of these species allows for the spin changes required for such type of reactions.

The enzymes able to activate molecular oxygen and incorporate both of its atoms in the aliphatic product are known as dioxygenases [53, 54, 55]. They are usually isolated from a variety of soil bacteria.
Ortho-hydroxyl dioxygenases are divided into two classes, according to the
the position at which the aromatic ring is cleaved, namely extradiol and intra-
diol dioxygenases. Extradiol dioxygenases cleave one of the C-C bonds adja-
cent to the hydroxyl groups of the substrate. They contain Fe(II) or Mn(II) as
a cofactor, coordinated by a 2-His-1-carboxylate facial triad [56], and solvent
molecules in square bipyramidal geometry. Intradiol dioxygenases cleave the
ring between the two adjacent hydroxyl substituents and utilize Fe(III) in their
active site. The iron is coordinated by two histidines, two tyrosines and a hy-
droxide. In both cases the organic substrate binds first to the metal ion, fol-
lowed by binding of dioxygen. A subsequent attack on the aromatic ring leads
to the bridging peroxide, which is believed to be common for the extradiol
and intradiol catechol dioxygenases. After O-O bond cleavage one of the oxy-
gen atoms is incorporated into the catechol ring via alkenyl migration in the
extradiol path, or acyl migration in the intradiol path (Scheme 6.1).

**Scheme 6.1:** Proposed catalytic cycles for extra- and intradiol cleaving dioxygenases.

This thesis presents theoretical studies of the reaction mechanisms of
three dioxygenase enzymes: Homoprotocatechuate 2,3-dioxygenase isolated
from *Brevibacterium fuscum*, Manganese-Dependent Homoprotocatechuate
2,3-Dioxygenase, and Homogentisate Dioxygenase.
**Scheme 6.2**: The overall reaction catalyzed by **A**: homoprotocatechuate dioxygenases *Bf* 2,3-HPCD and MndD; and **B**: Homogentisate Dioxygenase HGD.

*Bf* 2,3-HPCD and MndD are metalloenzymes that use molecular oxygen to cleave the aromatic ring of 3,4-dihydroxyphenylacetate, also known as homoprotocatechuate (HPCA), in an extradiol manner. Homogentisate dioxygenase is an enzyme involved in the catabolism of tyrosine and phenylalanine in human [57]. It catalyzes the oxidative ring scission of Homogentisate (HG) to maleylacetoacetate. The type of ring cleavage (between the carbons C1 and C2 substituted with hydroxyl and carboxymethyl groups, see Scheme 6.2 **B**) resembles the reaction catalyzed by Fe(III)-dependent intradiol dioxygenases. However, the dependence on Fe(II) and the structure of the active-site indicate that HGD is more related to the extradiol dioxygenases, which cleave the catechol ring at a bond adjacent to the carbons binding the hydroxyl groups. The reactions catalyzed by *Bf* 2,3-HPCD and MndD, and by HGD are summarized in Scheme 6.2 **A** and **B**, respectively.

The selectivity in the ring cleavage reaction, shown by extra- and intradiol dioxygenases is still not well understood. The metal ion requirement is an interesting issue – *Bf* 2,3-HPCD and MndD are highly identical in terms of amino acid sequence and they catalyze the same reaction, but they use a different metal as cofactor – Fe(II) and Mn(II), respectively. Studying these sys-
tems is important for improving our understanding of catalysis. Apart from that, HGD is interesting from a medical point of view. This enzyme is connected with two metabolic disorders – alkaptonuria (a rare hereditary human disease [58]) and tyrosinemia type I (a life-threatening disease caused by dysfunction of fumarylacetoacetase, an enzyme located downstream from HGD in the tyrosine catabolic pathway). In the studies presented in this part, hybrid DFT with the B3LYP functional was used with models of different size built from the crystal structures of the enzymes. The energetics of the suggested reaction path is investigated by locating intermediates and transition states and evaluating their relative energies, including environmental effects of the surrounding protein.
This section summarizes the results originally reported in Paper II and Paper III.

7.1 Background

*Bf* 2,3-HPCD and the manganese dependent analogue MndD have 83% sequence identity [59]. In both enzymes the active-site metal ion is coordinated by His155, His214 and Glu267 forming the 2-His-1-carboxylate facial triad occupying one face of an octahedron. This motif, typical for a number of non-heme Fe(II)-dependent enzymes involved in dioxygen activation [60, 61], is considered as a very important factor providing the right environment for the dioxygenation. The remaining three coordination positions of the octahedral complex are occupied by solvent molecules and are available for the substrates – Homoprotocatechuacte (HPCA) and dioxygen. High-resolution X-ray structures for both iron- and manganese-dependent homoprotocatechuacte 2,3-dioxygenases have been determined, and spectroscopic and steady state kinetic investigations have been performed, revealing important structural and mechanistic features of these biocatalysts. The activation barrier for both enzymes, predicted from the experimental data using transition state theory is about 16 kcal/mol [62, 63].

According to experimental data extradiol dioxygenases use an ordered mechanism for binding substrates. The aromatic substrate first binds, which activates the metal ion for O\textsubscript{2} binding. The O\textsubscript{2} adduct has not been detected and characterized up to date for the natural substrate. There is however evidence for NO binding [64, 65] supporting this scenario. The crystal structure of the enzyme with HPCA bound in the active-site (PDB code:1QOC; Figure 7.1 for *Bf* 2,3-HPCD) reveals that the substrate fills two of the free coordination sites, *trans* to the histidines, sites previously occupied by water ligands [59]. The remaining sixth coordination site is empty in *Bf* 2,3-HPCD, while it is occupied by a water ligand in MndD. This position is open for accommodation of O\textsubscript{2} later. HPCA binds to the metal center by its
two hydroxyl groups in an asymmetric manner, with a difference of about 0.2 Å between the two O-M(II) distances.

The acetate substituent of the substrate interacts with several residues in the active-site – it makes a salt bridge with the guanidinium group of Arg243 and is also forming a hydrogen bond with $N_D$ of His248. It was recently shown by mutagenesis studies of *Bf* 2,3-HPDC with the natural and with alternative substrates, that these interactions are quite important for the formation of the enzyme-substrate complex, and for the overall catalytic activity. Substrates without the carboxylate (4-methylcatechol) or the methylene group (3,4-dihydroxybenzoate) binds 30 and 555 times weaker, than the native substrate [62], respectively. It was even shown [63] that the His200Phe mutant of *Bf* 2,3-HPDC can perform *intradiol* cleavage of an alternative substrate 2,3-dihydroxybenzoate (2,3-DHB), where the side chain of HPCA is missing. This implies that the proper substrate binding is crucial for the position of the ring cleavage, since the native substrate is properly cleaved by the mutant, although with much lower activity.

Several proposals have been made about the catalytic mechanism of extradiol dioxygenases [59, 66, 67, 68, 69]. Scheme 7.1 summarizes one suggested proposal, in which the catecholic substrate binds to the metal cen-
7.1. Background

Scheme 7.1: The reaction mechanism proposed by Que et al. [59].

The reaction mechanism involves the following steps:

1. The substrate is deprotonated by His200, which assists in the deprotonation of the substrate.
2. The substrate becomes a semiquinone radical.
3. The substrate oxidation is facilitated by the removal of its second proton.
4. The bound O$_2$ is activated by transfer of one electron from the substrate through the metal, which becomes a superoxide (B in Scheme 7.1).
5. A nucleophilic attack of the superoxide on the electron-deficient substrate gives an alkyl-peroxy intermediate (C).
6. This intermediate undergoes a Criegee rearrangement consisting of O-O bond cleavage and insertion of one oxygen atom into the ring.
7. A lactone intermediate (AA) is generated and subsequently hydrolyzed by the hydroxide group remaining on the metal, leading to the α-hydroxy-δ-carboxymethyl $cis$-muconic semialdehyde product (Scheme 7.1).
8. Such a two-step oxygen insertion into the ring is supported by the partial exchange of $^{18}$O from O-labeling experiments [71].

Experimental studies providing strong support for the above mechanistic scenario were recently reported [72, 73]. Three intermediates in the catalytic cycle of $Bf$ 2,3-HPCD were detected, using an alternative
7.2. The model

Small models including only the most relevant parts of the residues directly coordinated to the metal center (His155, His214, Glu267 and Tyr257), and the truncated substrate molecule, were used initially to reduce the computational costs. Such models were however found insufficient for proper modeling of the enzymatic reaction. Although the role of His200 is not yet completely clear, this residue is undoubtedly important for the catalytic activity. Since it is assumed that the initial proton transfer from HPCA to the superoxide is assisted by His200, its inclusion in the model is logical. The electron transfer from the substrate to $O_2$ could be better reproduced using the whole substrate molecule, than with the truncated model (Paper II). Furthermore, it was found necessary to include His248 in protonated form, in order to maintain the neutral charge of the model. This residue also hydrogen-bonds with the carboxylates of Glu267 and HPCA and thus keeps the whole structure closer to the X-ray geometry. Asn157 was also added, since it provides some stabilization to the superoxide moiety. Thus a model consisting of 77 atoms was constructed and used for both $Bf\ 2,3$-HPCD and MndD. All the results given in the next sections were obtained with this model, unless the model is explicitly specified.

7.3 Reaction mechanism of $Bf\ 2,3$-HPCD

The calculated $Bf\ 2,3$-HPCD catalytic cycle is described in details in Paper III and basically follows the experimentally suggested one in Scheme 7.1. The HPCA substrate first binds at the active-site as a monoanion. Next the dioxygen molecule binds to the complex. The calculated binding energy of $O_2$ is $+2.9$ kcal/mol (solvent effects included), which would become around $+13$ kcal/mol if the entropy is added. A correction of about $-10$ kcal/mol to the QM results was applied and thus the estimate is that dioxygen is unbound by about $3$ kcal/mol. Three possible spin states were considered for the adduct with $O_2$ – triplet, quintet and septet. The septet state ($S=3$) was found to be the ground state, but the state of main interest was the quintet ($S=2$). In the quintet state the spin population is $3.75, 0.90$ and $-0.80$ for iron, dioxygen and the substrate, respectively, i.e. dioxygen is reduced with one electron to su-
peroxide \( \text{O}_2^{2-} \), and HPCA is oxidized and ready for further attack. The corresponding intermediate in Scheme 7.2 is denoted \(^5\text{R1}\). The energy of this state is spin corrected using the Heisenberg Hamiltonian approach. The correction is calculated to be 1.4 kcal/mol, putting the quintet intermediate 8.7 kcal/mol above the zero level. The energies of all the structures with open-shell low-spin coupling, reported furthermore, are corrected using this method. All the corrections are in the order of 1 kcal/mol and they do not change the relative energies or the conclusions.

The next intermediate is the bridging hydro-peroxide \( \text{BHP} \) formed by the attack of the activated dioxygen on the aromatic ring of the substrate. An approximate transition state \( \text{TS1} \) (Figure 7.2) for proton transfer from HPCA to dioxygen was found, leading to a \( \text{Fe(II)-O(H)-O} \) intermediate denoted \( \text{I1} \). Energetically \( \text{TS1} \) lies 17.4 kcal/mol above the zero level. \( \text{I1} \) is at 8.1 kcal/mol and the further attack on the aromatic ring yielding the hydro-peroxide \( \text{BHP} \) was found to have very small barrier, which even disappeared in the big basis set calculations. Direct formation of a bridging peroxide (\( \text{BP} \) in Figure 7.3 A) was also obtained. This intermediate is at an energy level of 15.3 kcal/mol (\( \text{BP} \) in Scheme 7.2). Going from \(^5\text{R1}\) to \( \text{BP} \) involves only the endothermicity of 6.6 kcal/mol, but no additional energy barrier. A transition state for proton
transfer from His200 to the peroxide was optimized (Figure 7.3 B) but the big basis set and dielectric effects put this transition state 0.7 kcal/mol below BP, thus leading to BHP without any activation barrier (Scheme 7.2). Considering the accuracy of the method it can be concluded that these two paths from $^5\text{R1}$ to BHP are equally probable since they involve basically the same activation barrier. The $^5\text{R1}$-to-BHP reaction step is a good candidate for the rate limiting step for the whole catalytic cycle, since its calculated activation barrier is in very good agreement with the transition state theory prediction from the experimental rate of 16.1 kcal/mol.

The O-O bond cleavage of BHP goes through a transition state TS2, for which the spin population on Fe is $S_{Fe}=3.94$. This corresponds to oxidation state between Fe(II) and Fe(III), i.e the metal ion is oxidized. The spin population on the oxygen atom bound to the substrate is -0.40, which clearly shows that the cleavage is homolytic. This step requires activation energy of 11.3 kcal/mol, and leads to an epoxide intermediate EE through a metastable (transient) radical state denoted RS. Then the C-C bond of the ring is cleaved with an activation barrier of 8.0 kcal/mol (TS3 in Scheme 7.2). Thus the suggested Criegee rearrangement was found to occur stepwise instead of concerted. As a result a lactone intermediate LR is formed, which is then attacked by the OH group bound to the Fe(III) center. Finally the ring of the resulting semialdehyde LR-OH opens to give the reaction product.
7.4 Reaction mechanism of MndD. Spin transition

The energy profile shown in Figure 7.3 was obtained in the initial study of MndD, described in Paper II. The calculated catalytic cycle for the sextet spin state follows the experimentally suggested path (Scheme 7.1) as in the case of Bj 2,3-HPCD. In summary, the HPCA substrate first binds at the active-site as a monoanion, followed by dioxygen. The latter is activated to a superoxo radical by an electron flow from the substrate through the metal center, and reacts with HPCA forming the bridging hydro-peroxide structure BHP at the C3 position of the ring (Figure 7.5 A). Importantly, this intermediate can be formed only after the preceding protonation of O2. In the calculations this protonation is modeled by a proton transfer from the C4 bound hydroxylic group of HPCA, assisted by the second shell His200, i.e. the complete deprotonation of the substrate is reached. It should be noted here that based on recent mutagenesis experiments, reported by Emerson et al. [70], it was argued that His200 mainly stabilizes O2 in the enzyme-substrate-dioxygen complex. Both possibilities, that His200 can act as a base to deprotonate the second hydroxide group of the substrate, or as an acid protonating the reduced oxygen were found unlikely by Emerson’s study. Those results show that the key role of the residue in this position is related to its ability for making hydrogen bonds, since the mutants having His200 replaced by asparagine or glutamine preserve 60-80% of the native activity. The calculations rather show that His200 can facilitate the proton transfer and can further stabilize the Mn(II) bound hydroperoxide. In that sense, there is no conflict between the computational results and the experimental findings, because there is no net change of the protonation state of His200 in the proton shift.
7.4. Reaction mechanism of MndD. Spin transition

Figure 7.4: Optimized structures of transition state for O-O bond cleavage TS2. The most important distances and spin populations (in the text-box) are given.

Scheme 7.3: Energetic profile of the suggested catalytic pathway for the sextet spin state of MndD.

The possibility for an attack of the peroxo species at the C4 position of the HPCA aromatic ring was also considered. The results are described in details.
in Paper II. A bridging peroxide at the C4 position can be formed before, as well as after the preliminary protonation of the superoxo radical. These alternative reaction paths however either involve higher activation barriers, or lead to unstable products, compared with the C3-attack and formation of BHP.

Figure 7.5: Optimized structures of A: the bridging hydro-peroxide BHP; B: the transition state for O-O bond cleavage. The most important spin populations (bold) and distances (italics) are given.

The next step – the homolytic O-O bond cleavage was found to be the rate limiting step in the suggested mechanism, with an activation barrier of 21.8 kcal/mol (Scheme 7.3). It leads to a transient radical intermediate RS, which easily collapses to an extra epoxide structure EE going 13.9 kcal/mol down in energy. At this stage, the competitive intra cleavage reaction was found to be possible. Thus, the calculated barrier for intradiol cleavage is 4.5 kcal/mol (see TSi in Scheme 7.3), appearing to determine the selectivity between both reactions. However, some uncertainty about the origin of this difference remains, since the removal of His248 or Asn157 separately does not change the picture, while their simultaneous removal leads to the disappearance of the intra cleavage barrier. The C-C bond cleavage of the extra epoxide EE goes over a barrier of 7.3 kcal/mol, leading to the lactone ring intermediate LR with 13.3 kcal/mol exothermicity. Next, the Mn(III) bound hydroxyl group attacks the lactone ring without any significant barrier, going another 18.1 kcal/mol down in energy to the intermediate denoted as LR-OH in Scheme 7.3. The final steps of the reaction including the ring opening of the LR-OH intermediate and the product formation do not involve any significant barriers and were not studied in details.

The above results were obtained following the potential energy surface of the sextet spin state, which is the ground state of the resting enzyme. The
relatively high calculated activation barrier of 21.8 kcal/mol with respect to the experimentally observed 16.3 kcal/mol, was initially assigned to the usual tendency of the B3LYP method to slightly overestimate the reaction barriers. The results obtained for \textit{Bf} 2,3-HP, however, clearly suggest that additional investigation should be carried out for the MndD enzyme. These studies are reported in \textbf{Paper III}. The attention was focused on the O-O bond cleavage step, where a large difference between the Fe- and Mn-containing enzymes of almost 10 kcal/mol was obtained. A large number of calculations with different models were performed attempting to clarify the origin of this difference. The main result is that the difference of about 10 kcal/mol between \textit{Bf} 2,3-HP and MndD, regarding the activation barrier for O-O bond cleavage, was reproduced with all models probed. No significant geometrical difference was found in any model between the corresponding intermediates and transition states of the two enzymes. The only difference between the two cases was the coupling of the spin of the radical substrate (\textit{subO}*) with the spin of the metal ion. In the \textit{Bf} 2,3-HPD case this spin was always antiferromagnetically coupled to the unpaired electrons of iron, while in MndD this coupling was always ferromagnetic (Table 7.1).

<table>
<thead>
<tr>
<th>spin</th>
<th>(S_M)</th>
<th>(S_{\text{subO}^*})</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS2_{\text{Bf,3-HPD}}</td>
<td>3.94</td>
<td>-0.40</td>
</tr>
<tr>
<td>TS2_{\text{MndD}}</td>
<td>4.40</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 7.1: \textit{Spin populations on the metal ion M and subO* in the corresponding O-O bond cleavage transition states found for Bf 2,3-HPD and MndD.}

Antiferromagnetic coupling in MndD is possible for the quartet spin state (S=3/2). Therefore the quartet state potential energy surface for MndD was investigated for the O-O bond cleavage from the BHP reactant to the EE product. The quartet BHP is 24.2 kcal/mol higher in energy than the corresponding sextet structure, while the quartet EE is as stable as the sextet product. Thus it was assumed that the activation barrier for the homolytic O-O bond cleavage in MndD could actually be determined by a spin transition from the sextet to the quartet state, occurring at lower energy than the sextet transition state. Such spin transition should occur in a region where the two spin surfaces cross. A crossing point slightly before the sextet transition state (Figure 7.6 A) was found in a rough scan along the O-O bond on both PESs. The \textit{lacvp} energy of this point is around 16 kcal/mol (relative to the sextet \(^6\text{BHP}\)). With this approximate crossing as a starting point, two interatomic distances were chosen as reaction coordinates, namely the O-O and Mn-O distances, and a two-dimensional potential energy surface scan was done. These two distances
7.4. Reaction mechanism of MndD. Spin transition

Figure 7.6: A potential energy diagram for the O-O bond cleavage step in MndD; B qualitative picture of the sextet and quintet PESs; the plots are generated on the basis of 110 grid points obtained in the geometry scan for each surface.

were scanned from 1.5 to 2.5Å and from 2.5 to 1.7Å, respectively, with 0.1Å step size. As a result a grid of points was obtained for each PES, and the surfaces were plotted. Figure 7.6 B gives a qualitative picture for the behavior of the two surfaces. The quartet surface is definitely repulsive towards the reactant BHP and goes directly downhill to the product region. On the sextet surface one can clearly see both stationary points and the minimum-energy path connecting them with the transition state corresponding to TS2. The scan data shows that the two states have closest energies at an O-O distance of 1.8Å and a Mn-O distance of 2.0Å. In order to determine more accurately the spin
transmission geometry a higher resolution geometry scan (step size 0.02Å) was performed around this point. Big basis set calculations were carried out for each geometry for obtaining more accurate energies. As a result a point was found with an O-O distance of 1.80Å and a Mn-O distance of 2.04Å, and with an energy of 14.4 kcal/mol relative to the reactant BHP. Attempts to find the Minimum-Energy Crossing Point (MECP) [74] were made using a separate program [75]. Since this program looks not just for any crossing point between two surfaces, but for the minimum-energy crossing point, the search actually ended up in the next crossing point, where the system comes back to the sextet spin state. This point turned out to be the EE epoxide product of the cleavage (Figure 7.6A).

![Scheme 7.4: Updated potential energy diagram for MndD.](image)

Thus, it can be concluded that the O-O bond cleavage in MndD involves an activation barrier of 14.4 kcal/mol, due to a spin transition between the sextet and the quartet spin states. The spin-orbit coupling of the Mn ion makes this transition possible. Scheme 7.4 shows the updated energy diagram of the calculated reaction mechanism.

It can be summarized that according to the calculations Bf 2,3-HPCD and MndD follow reaction paths involving the same mechanistic steps, but with slightly different energetics. The overall rate-limiting step for Bf 2,3-HPCA
7.4. Reaction mechanism of MndD. Spin transition

is the formation of the alkylperoxo intermediate **BHP**. For MndD a spin transition from the sextet to the quartet spin state after **BHP** was found to determine a barrier of 14.4 kcal/mol for the homolytic O-O bond cleavage. Thus the experimentally observed similarities between Bf 2,3-HPCD and MndD are reproduced in the calculations, and some important differences could be determined.
8. Homogentisate Dioxygenase

This section summarizes the results originally reported in Paper I.

8.1 Background

Despite the considerable importance of HGD, only recently was some progress made in the research on this interesting enzyme. The crystal structure of human HGD was solved by Titus and co-workers [76]. As in other extradiol dioxygenases, the active-site iron is coordinated by two histidines (His335, His371) and one glutamate (Glu341), which form the 2-His-1-carboxylate facial triad. The other three coordination positions around iron are available for Homogentisate and dioxygen or solvent ligands. Several other residues were proposed to be involved in substrate binding and catalysis. First, His365 forms a hydrogen bond with Glu341 and may be involved in substrate deprotonation upon binding. Second, His292 was proposed to form a hydrogen bond with the C4-bound hydroxyl group of HG, and together with Pro295, Pro332, and Tyr333 forms a solvent-accessible box, which was also implicated for the substrate binding.

Scheme 8.1: Reaction Mechanism proposed by Titus et al. for HGD.
Recently published kinetic analysis has provided firm evidence for an ordered mechanism, in which HG binds first to the enzyme followed by a dioxygen reacting with the Fe(II)-HGD-HG complex \[77, 78\]. Furthermore, an optimum pH (6.2) and a turnover number (16 \(s^{-1}\)) were measured. Aside from this, little is known about the catalytic mechanism of HGD. The mechanism presented in Scheme 8.1 was proposed \[76\], and with the exception for the chemical identity of the substrate and the involvement of His292, it is very similar to the mechanism proposed by Lipscomb and co-workers for extradiol dioxygenases \[65, 69\]. Thus, it is proposed that HG chelates Fe(II) with its carboxyl and deprotonated hydroxyl groups. Molecular dioxygen reacts with this HGD-HG complex and forms a peroxo intermediate with an O-O bridge joining the C2 carbon and Fe(II). In the next step, the peroxo intermediate undergoes a Criegee rearrangement, yielding a seven-membered lactone ring, protonated His292, and a hydroxyl group bound to Fe(II). Finally, the lactone ring is hydrolyzed to an enol form of the product, maleylacetoacetate. There are experimental findings supporting the formation of the lactone intermediate during the extradiol ring cleavage \[71\]. However, there are also some recent experimental data which are difficult to explain with such a generic mechanism for Fe(II)-dependent extradiol dioxygenases. More specifically, the discovery of salicylate 1,2-dioxygenase and the observed stereo-chemistry of its product (\(trans\) configuration around the C3-C4 bond) show that the substrate does not have to possess the second hydroxyl group (or its equivalent) and that the ring scission mechanism is probably more complicated, at least for this particular dioxygenase \[79\].

8.2 The model

Available crystal structures for HGD were solved for the apo and holo forms of the enzyme, i.e there is currently no structure for the substrate-enzyme complex. For this reason, the first task was to investigate the structure of the HGD-HG complex. A large model of 134 atoms was used, comprising the iron first-shell ligands (His335, His371, Glu341, and HG), two second-shell histidines (His292 and His365), and three groups defining a substrate box (Pro295, Pro332, and Tyr333). Appropriate models are used for the different amino acid residues included. HG is assumed to be a dianion, His365 is doubly protonated (cation), while His292 is neutral and accepts a hydrogen bond from the phenolic group of HG. The total charge of the model is zero. Smaller model (80 atoms) was used to study the reaction mechanism.
8.3 Enzyme-Substrate Complex

Homogentisate binds to the active-site iron through its carboxyl group and the deprotonated C1-bound hydroxyl functionality [76]. Two arrangements around iron are possible. In the first one, the position trans to Glu341 is vacant, and HG occupies sites trans to His335 and His371 (Figure 8.1 A). In the second arrangement, the vacancy is trans to His371, and HG is bound at sites trans to Glu341 and His335 (Figure 8.1 B). The calculated energy difference for these two optimized models is 5.2 kcal/mol in favor of the first one (A). This energy splitting remains almost unchanged when the size of the model is reduced (4.9 kcal/mol calculated for the smaller model), therefore it is concluded that the difference comes from the arrangement of the ligands around iron.

8.4 Reaction mechanism of Homogentisate Dioxygenase

The computational results suggests that the catalytic reaction of HGD involves three major steps (Scheme 8.2). First, molecular dioxygen binds to iron and reacts with homogentisate, forming the peroxo-bridged intermediate (1, 2, and 3). Dioxygen binds to the active-site iron trans to Glu341. The ground state of the resulting ternary complex (HGD-HG-O2) is a septet, which is 1.1
kcal/mol above the zero level corresponding to separate reactants. The reactive state however is the quintet $^5$1 which is only 3.0 kcal/mol above the septet (Scheme 8.3). The attack of the dioxygen fragment on the C2 carbon of the HG ring leads through the transition state TS1 (Figure 8.2) with an activation barrier of 18.9 kcal/mol. This is apparently the highest barrier along the suggested reaction path. The initial attack of the Fe-bound O$_2$ fragment on the aromatic ring results in a peroxy-bridged structure 2. Indeed, the ground state for the peroxy-bridged structure is species $2a$, which lies 6.4 kcal/mol above the zero level. In this structure, the proton has been moved from His365 to the carboxylate group of the substrate, and this protonated carboxylate forms a strong hydrogen bond with the distal oxygen of the peroxy group (Figure 8.3). The direct proton shift from the carboxyl to the peroxy group yields intermediate 3, which is 3.1 kcal/mol less stable than $2a$ (Scheme 8.3).

The second major step is a homolytic cleavage of the O-O bond in the hydroperoxo intermediate 3, leading to the Fe(III)-OH/alkoxyl radical species 4. The calculated activation barrier is 10.7 kcal/mol, and the reaction ($2a \rightarrow 4$) is exothermic by 1.6 kcal/mol. The O-O bond cleavage is promoted by an electron transfer from Fe(II) to the peroxy group. The product of the O-O bond homolysis, i.e. the alkoxyl radical species 4, has been found to be a meta-stable species, since it collapses to the arene oxide radical intermediate 5 without any activation barrier.
8.4. Reaction mechanism of Homogentisate Dioxygenase

Scheme 8.3: Calculated energy profile along the suggested reaction path for HGD.

Figure 8.2: Optimized structure **TS1** for transition state for an attack of the dioxygen ligand at the aromatic ring. Distances in Å in bold, spin populations in italics.

These two initial steps are similar to the corresponding reactions proposed for *Bf* 2,3-HPCD and MndD. Alternative reaction mechanism which leads to the bridging peroxo structure has been investigated. Taken together, there are at least three pathways which involve comparable barriers. Discussion for the less likely routes can be found in [Paper I](#).
Figure 8.3: Optimized structure 2a for peroxo-bridged intermediate with a hydrogen bond between protonated carboxyl and peroxo groups. Distances in Å in bold, spin populations in italics.

The next phase of the catalytic cycle involves a nucleophilic attack of the Fe-bound OH group on the carbonyl carbon of the arene oxide radical (TS4 in Scheme 8.2). The calculated activation barrier is only 6.5 kcal/mol, and the reaction is endothermic by 4.3 kcal/mol (Scheme 8.3). From the product 7 the epoxide ring opens through the meta-stable structure 9, and the following transition state TS6. Concerning the energetics, the relative energy of TS6 is -7.4 kcal/mol, which means that the accumulated barrier calculated with respect to the arene oxide radical 5 is 12.3 kcal/mol. The spin populations of the product 10 indicates that the ring has been cleaved to the closed-shell enolic form of maleylacetoacetate, while the active-site iron is back in the high-spin Fe(II) electronic state.

The alternative path (TS3, 6, and TS7) and the plausible side reaction (TS5 and 8) involve markedly higher barriers than those of the competing reactions of the proposed mechanism (Scheme 8.3).
9. Summary

All three enzymes bind the substrates in an ordered manner. After reducing the dioxygen by electron transfer from the organic substrate, a peroxo-bridged intermediate is formed. This reaction step is found to be the rate-limiting step for the two Fe(II)-dependent enzymes *Bf* 2,3-HPCD and HGD. The second major step is the homolytic cleavage of the O-O bond in the peroxo intermediate. Notably, for the Mn(II)-dependent enzyme this cleavage happens on the crossing between the quartet and the sextet spin states, and defines the rate limiting activation barrier.
Part IV:
Biomimetic model complexes

Two synthetic Fe complexes were studied computationally using Density Functional Theory. The first one is a non-heme \([(L-N_4Me_2)Fe(III)]^{+3}\) complex that mimics the reactivity of intradiol dioxygenases. The second one is a heme \([T(o-Cl)PPFe]\) complex, which catalyzes the stepwise oxidation of cyclohexane to adipic acid.
10. Non-heme iron model complex for intradiol-cleaving dioxygenases

This chapter summarizes the results originally reported in Paper IV.

10.1 Background

Two Fe(III)-catecholate complexes have been synthesized and their crystal structures determined [80]. In these complexes, iron is coordinated by a tetraazamacrocyclic ligand L-N₄Me₂, with two adjacent coordination sites available for the binding of the substrate. It has been experimentally shown that the complexes react with molecular oxygen at room temperature and exhibit catalytic activity for intradiol-cleavage of the substrate. The experimentally determined velocity constants for the substrates catechol and ditertbutyl catechol correspond to activation barriers of 21.0 and 18.0 kcal/mol, respectively. In comparison for intradiol cleaving Catechol Dioxygenase [81] the rate constant \( k = 2.5 \times 10^5 \text{ M}^{-1}\text{s}^{-1} \) corresponds to a barrier of 10.1 kcal/mol. The basic chemical steps of the intradiol dioxygenation are the following: (i) binding of dioxygen to iron, accompanied by oxidation of the substrate to semiquinone and reduction of dioxygen to superoxide, (ii) formation of peroxo-bridging intermediate after an attack of the superoxide on the semiquinone ring, (iii) O-O bond cleavage assisted by protonation of the peroxide, (iv) acyl migration leading to an anhydride intermediate, which is then hydrolyzed to the acyclic product. Even though the coordination environment of the model complex and the enzyme are quite different, questions regarding the C-C cleavage step, the actual attack of the dioxygen molecule on the Fe(III)-catechol system and the intradiol specificity can be addressed.

10.2 The model

The computational model used in this work is based on the crystal structure reported from Koch an co-workers [80]. In terms of computational procedure and methodology the study was done according to the descriptions in Chapter 5 of Part 2, with few differences compared to the enzymatic investigations. Together with the original hybrid density functional B3LYP, B3LYP∗
was used, for which the amount of Hartree-Fock exchange is reduced from 20% to 15% [33, 82]. In this study the rate limiting step was lowered by 5.4 kcal/mol when using B3LYP* instead of B3LYP.

Zero-point effects, entropic and thermal corrections to the Gibbs free energy were evaluated for each stationary point. It is difficult to describe correctly the entropic contributions for the oxygen binding step within the present modeling. The calculated entropic contribution for the oxygen binding is 8.6 kcal/mol, which goes up to 13.4 kcal/mol during the formation of the bridging peroxide. Once dioxygen is bound, and the reaction has started, the relative entropy effects vary by 1 to 2 kcal/mol. To account for compensating effects in the oxygen binding a Van der Waals correction was evaluated with B3LYP-D using the Orca Package [83, 84]. Corrections in the range of 3 to 4 kcal/mol were calculated for these long-range dispersion interactions. The solvent was modeled as a macroscopic continuum with a dielectric constant $\varepsilon=36.6$ (corresponding to acetonitrile). Final energies of the optimized structures were corrected for the solvent effects by employing the $l_{acyp}$ basis set. As already mentioned, the dielectric medium should have a small effect on the reaction energetics as long as no substantial charge separations are involved. However, some charge separation is present in this case and solvent effects therefore range from 4 to 10 kcal/mol. Reported spin populations are used to indicate the spin and the oxidation state of the metal ion and were derived from a Mulliken population analysis.

![Scheme 10.1](image1)

Scheme 10.1: Proposed mechanism for intradiol cleavage catalyzed by the Fe(III)L-N$_4$Me$_2$ complex (where L-N$_4$Me$_2$ is N,N'-Dimethyl-2,11-diaza[3.3](2,6)pyridinophane) based on the calculations.
10.3 Reaction mechanism

The proposed mechanism based on our calculations is given in Figure 10.1. Because of the open-shell nature of the transition metal the starting Fe(III) complex (denoted as \( \textbf{R1} \)) can have three different spin states – sextet (spin \( 5/2 \)), quartet (spin \( 3/2 \)) and doublet (spin \( 1/2 \)). Due to possible reduction of the iron by the substrate a total number of 8 spin states exist for \( \textbf{R1} \). This number originates from various couplings between the unpaired electrons arising upon one electron transfer from the substrate to Fe. In order to explore the reaction coordinate one needs to know the energetic splitting of these states for all the reaction intermediates. Some of these states were found to be energetically high and thus thermally inaccessible. A summary of the calculated relative energies for the lowest spin states for the reactant is given in Table 10.1.

<table>
<thead>
<tr>
<th>S</th>
<th>ox. state</th>
<th>Fe</th>
<th>catechol</th>
<th>( \Delta G )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>Fe(II)</td>
<td>-</td>
<td>1( \alpha )</td>
<td>0.0 (3.8)</td>
</tr>
<tr>
<td>3/2</td>
<td>Fe(II)</td>
<td>2( \alpha )</td>
<td>1( \alpha )</td>
<td>0.8 (3.3)</td>
</tr>
<tr>
<td></td>
<td>Fe(II)</td>
<td>4( \alpha )</td>
<td>1( \beta )</td>
<td>4.5 (1.5)</td>
</tr>
<tr>
<td></td>
<td>Fe(III)</td>
<td>3( \alpha )</td>
<td>-</td>
<td>8.8 (10.8)</td>
</tr>
<tr>
<td>5/2</td>
<td>Fe(II)</td>
<td>4( \alpha )</td>
<td>1( \alpha )</td>
<td>1.2 (0.0)</td>
</tr>
</tbody>
</table>

Table 10.1: *Spin-splitting in the reactant \( \textbf{R1} \). Relative free energies given in kcal/mol. The values in parenthesis are calculated using the standard B3LYP functional.*

The ground state of the reactant predicted by the calculations is a doublet state corresponding to low-spin Fe(II) and one electron oxidized catecholate radical (semiquinonate), i.e. \( [(\text{LN}_4\text{Me}_2)\text{Fe(II)(cat)}^-]^{+} \), (Figure 10.1 A). A quartet state 2\( \alpha 1\alpha \) formed by intermediate spin on Fe(II) and semiquinonate is only 0.8 kcal/mol up in energy, and a sextet 4\( \alpha 1\alpha \), corresponding to a high-spin Fe(II) ferromagnetically coupled to the one electron oxidized semiquinonate lies at +1.2 kcal/mol. There are two more accessible quartet states. One is the antiferromagnetically coupled reactant Fe(II)-\( \text{cat}^{\cdot -} \), 4\( \alpha 1\beta \), at 4.5 kcal/mol above the ground state. The other quartet is formed by intermediate spin on Fe(III) and catecholate \( \text{cat}^{2-} \) and lies at +8.8 kcal/mol. However, no high spin Fe(III) catecholate complex could be found. Thus, states having an oxidized monoanionic substrate with a reduced Fe(II) ion are preferred for the reactant structure. As can be seen from Table 10.1 B3LYP* changes the spin-splitting, stabilizing the low-spin, and destabilizing the high-spin states. The dominating ferrous-semiquinonate resonance is supposed to direct the dioxygen binding directly to the metal. In this respect
the computational results for the biomimetic model complex considered here are in line with previous theoretical [85, 86] and experimental [7] findings concerning intradiol cleavage in enzymes.

Figure 10.1: Optimized structures of A: The reactant R1 in its ground state; B: the superoxide complex R2.

The reaction starts with formation of the oxygen adduct R2 from R1 and $^3$O$_2$. This leads to one electron reduction of the dioxygen molecule to a superoxide O$_2^\bullet^-$. The electron is provided by the substrate radical, which in turn becomes a neutral quinone. One oxygen atom of the substrate has to dissociate from the metal ion upon oxygen binding to provide the necessary coordination site, see Figure 10.1 B. Two spin states for the R2 complex with quite close energies were identified, a quartet $^4R_2$ and a sextet $^6R_2$, where the unpaired electrons are residing on iron and oxygen, respectively (no spin on the substrate). The quartet state lies at 18.0 kcal/mol and the sextet state at 18.2 kcal/mol relative to the reaction zero level. Additionally, a doublet state $^2R_1$ exists at an energy level of 26.6 kcal/mol.

The oxygen binding was investigated by using a test model including a protonated triethylamine [NH(Et)$_3$]$^+$, for N(Et)$_3$ is a base present in the experimental setup [80]. The result showed that most of the energy needed to form the oxygen bound adduct $^4R_2$ is due to the dissociation of the semiquinonate (details can be found in Paper IV). Direct dioxygen binding to the substrate instead of the metal was not possible to obtain without constraining the distance between a catecholate carbon atom and O$_2$. The resulting computational model complexes were found to be 8 to 12 kcal/mol higher than the corresponding metal-bound structures. It can be concluded that direct oxygen binding to the substrate is unlikely to occur. These results are in good agreement with the predominant ferrous character that was found in the R1 complex, which enables direct binding of the dioxygen to the metal.
In the next step the superoxide $\text{R2}$ attacks the aromatic ring of the substrate. With Fe(II) providing the second electron required for the reduction of the superoxide, a peroxide bridge is formed, and the metal ion is oxidized to Fe(III). The transition state $^4\text{TS1}$ (Figure 10.2 A) for this step lies at 23.3 kcal/mol on the quartet surface. No corresponding transition state was found on the sextet surface, instead, a transition state was found leading to another intermediate (see below). The formation of the bridging peroxide $^4\text{BP}$ (Figure 10.2 B) is slightly exergonic with a reaction free energy of 4.5 kcal/mol. The ground state for this intermediate is a quartet, with a doublet state around 10 kcal/mol higher energy. No bridging peroxide could be found on the sextet PES.

![Figure 10.2](image1.png)

Figure 10.2: Optimized structures of A: transition state for peroxide formation $\text{TS1}$; B: the bridging peroxide $\text{BP}$.

![Figure 10.3](image2.png)

Figure 10.3: Optimized structures of A: transition state for the homolytic O-O bond cleavage $^4\text{TS2}$; B: the alkoxy radical $\text{RAD}$, formed after the O-O bond cleavage.
From $^4$BP the reaction proceeds with a homolytic O-O bond cleavage resulting in formation of a radical state $^4$RAD (Figure 10.3 B), an Fe(IV)-oxo species containing an alkoxy radical on the substrate. The activation free energy for the bond cleavage is 8.0 kcal/mol with respect to $^4$BP, and the transition state $^4$TS2 is shown on Figure 10.3 A. The reaction free energy for this step is only 0.6 kcal/mol, i.e. the O-O bond cleavage is almost thermoneutral. The same kind of alkoxy radical intermediate was also found in previous studies on extra- and intradiol dioxygenases (Fe- and Mn-dependent) [87, 86] (Paper II and Paper III) and homogentisate dioxygenase (Paper I). In all these systems this short-lived radical intermediate spontaneously converts into the next intermediate. Such an alkoxy radical intermediate is absent in the case of a concerted Criegee rearrangement (a heterolytic cleavage process). The existence of the radical intermediate opens the possibility for the alternative extradiol reaction. The results describing the selectivity of the ring cleavage and thus the specificity of intra- versus extradiol ring cleavage are presented in the next section.

On the sextet PES the peroxide intermediate was not found. The sextet state $^6$R2 undergoes a direct homolytic O-O bond cleavage leading to Fe(IV)-oxo/alkoxy radical $^6$RAD. A transition state for this process was found and fully optimized, but it lies 15.3 kcal/mol above the corresponding $^4$TS2. Also, the resulting alkoxy radical intermediate is 7.4 kcal/mol less stable than the equivalent quartet state structure. These energy differences clearly indicate

Scheme 10.2: Calculated energy profile along the reaction coordinate; red – quartet state PES, blue – sextet state PES, solid line – the observed intradiol reaction path, dashed line – the alternative extradiol reaction path.
that the quartet path, going through a peroxo intermediate, is the preferred one for the reaction.

10.4 Intra- vs. Extradiol bond cleavage

From the Fe(IV)-oxo/alkoxy radical intermediate both intra- and extradiol cleavage reactions can occur. It is therefore important to be established which pathway is energetically favorable. A summary of the reaction profiles is shown in Figure 10.2. From \textbf{RAD} the intradiol path proceeds in a single step involving C-C bond cleavage and C-O bond formation. The transition state $^4\text{TSi}$ for this intradiol cleavage could be located on the quartet PES using the small basis set (Figure 10.4 \textbf{A}). However, the large basis set correction to the energy lowered the energetic position of that transition state below that of the radical state $^4\text{RAD}$. Therefore it can be concluded that within the computational accuracy, the formation of the intradiol product proceeds to \textbf{ANH} (Figure 10.5 \textbf{A}) with at least a small barrier.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig10_4.png}
\caption{Optimized structures of \textbf{A}: transition state for intradiol cleavage $^4\text{TSi}$; \textbf{B} transition state for epoxide formation $^4\text{TSe}$.}
\end{figure}

The reaction free energy obtained for the entire reaction is -33.9 kcal/mol. The sextet product $^6\text{ANH}$ was calculated to be 3.2 kcal/mol more stable than the quartet, which indicates that the system undergoes an inter-system crossing back to the resting sextet spin state.

The alternative reaction leading to an extradiol cleavage product proceeds via formation of an epoxide intermediate. On the quartet surface this requires an energy of 4.1 kcal/mol to overcome the activation barrier. The corresponding transition state is shown in Figure 10.4 \textbf{B}. For the sextet surface the barrier disappears when applying the big basis set correction. The differentiation be-
10.5. Comparison with the biological system

It is interesting to compare the mechanistic and energetic aspects of the catalytic reaction of the present biomimetic complex with the computational results for an intradiol dioxygenase Protocatechuate 3,4-dioxygenase (3,4-PCD) [86]. The mechanisms are very similar, involving essentially the same main chemical steps. Formation of a bridging peroxo intermediate after binding and one-electron reducing of the dioxygen, and cleavage of the O-O bond followed by insertion of the oxygen atom into the ring, are the key features of the intradiol aromatic ring cleavage. The calculated rate limiting step for the synthetic complex is the bridging peroxide formation, with activation barrier of 23.3 kcal/mol (for the quartet spin state), while for the enzymatic case it

Figure 10.5: Optimized structures of A: the anhydride product of the intradiol cleavage ANH; B: the lactate product of the extradiol pathway LA.
is the O-O bond breaking that determines the overall reaction rate with 20.8 kcal/mol activation barrier. The formation of the Fe(II)-(O_2)^•− adduct appears to be quite endergonic for the biomimetic complex. The reason is that the binding of the O_2 molecule requires dissociation of one of the substrate oxygens from iron, and opening of a free coordination site. Within the present model this transformation was found very expensive in terms of free energy. The enzymatic environment on the other hand, provides a coordination site for dioxygen in the beginning of the reaction, and facilitates later conformational change by stabilization interactions with second-shell residues. For the biomimetic system, the calculated sextet spin state reaction involves too high energies, but spin transition to a quartet state was found to occur, where the calculated activation free energy is in a good agreement with the experimentally determined one. A Criegee rearrangement transition state was found for the enzyme, which requires in-plane orientation of the two oxygen atoms and the two carbon atoms involved in bond cleavage and formation. This was not possible for the present model system, since the rigid ligand on iron does not provide enough flexibility for the proper arrangement.

From the calculated spin splittings it can be concluded that only the quartet state is likely to be involved in an energetically plausible mechanism. The sextet, being the resting state for the reactant and the product, is involved in the beginning and the end of the reaction. Doublet states were always found to be 5 to 15 kcal/mol higher in energy than the corresponding quartet and sextet intermediates (Table 10.2).

<table>
<thead>
<tr>
<th>intermediate</th>
<th>doublet</th>
<th>quartet</th>
<th>sextet</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2</td>
<td>26.6 (32.8)</td>
<td>18.0 (18.6)</td>
<td>18.2 (19.4)</td>
</tr>
<tr>
<td>BP</td>
<td>21.2 (27.7)</td>
<td>13.5 (18.5)</td>
<td>-</td>
</tr>
<tr>
<td>RAD</td>
<td>22.1 (30.6)</td>
<td>14.1 (20.7)</td>
<td>21.5 (27.6)</td>
</tr>
</tbody>
</table>

Table 10.2: Comparison of the relative energies (kcal/mol) for different spin states of intermediates along the reaction coordinate. The values in parenthesis are calculated using the standard B3LYP functional.

The calculated activation free energy of 23.3 kcal/mol for the rate limiting step is in a good agreement with the experimental barrier of 21.0 kcal/mol. The involvement of a peroxo intermediate in both intra- and extradiol cleavage reactions is in agreement with recent both computational and experimen-
tal studies. The peroxide O-O bond cleaves homolytically to give a radical Fe(IV)-oxo intermediate \textbf{RAD}. The ring cleavage and oxygen atom insertion to give the intradiol product proceed without an activation barrier. A muconic anhydride is formed, which is the final product within the scope of the present work. The alternative path of extradiol cleavage needs to cross an additional barrier that leads to the formation of an intermediate epoxide. Solely due to this barrier intradiol cleavage is observed for this biomimetic model complex.
11. Heme complex catalyzing adipic acid synthesis

11.1 Background

Adipic acid is the most important dicarboxylic acid for the industry, with main application as a precursor for the production of nylon and other polymers (PVC, Polyurethane). It is used also as a flavorant, acidulant and gelling aid in baking powders [88]. Since adipic acid occurs rarely in nature, it is synthesized in enormous amounts by the chemical industry. Historically, adipic acid was produced by oxidation of various fats, while the contemporary methods are based on oxidation of cyclohexane. First the cyclohexane is oxidized at 150-160° C with oxygen in the presence of homogeneous cobalt or manganese catalyst to a mixture of cyclohexanol and cyclohexanone called "KA oil" (abbreviation of "ketone-alcohol oil"). The KA oils can be further oxidized with nitric acid, hydrogen peroxide or molecular oxygen to adipic acid. The use of nitric acid has disadvantages from environmental point of view, but is preferred due the low cost of the nitric acid as oxidant. Hence, there is increasing interest towards usage of molecular oxygen or hydrogen peroxide, which provide far more environmental friendly oxidation.

The one-pot oxidation method has been reported from different groups [89, 90, 91]. The one-pot synthesis is a strategy where a reactant is subjected to successive chemical reactions in just one reactor. This improves the efficiency of a chemical reaction because a lengthy separation process and purification of the intermediates are avoided, while the yields are increased. This method however, has some disadvantages that still make the process expensive – use of expensive radical catalyst, acidic solvent, severe reaction conditions. Utilizing metalloporphyrins as catalysts for the one-pot synthesis of adipic acid from cyclohexane and molecular oxygen has been reported [91, 92] as promising, but up to date it is still hard to be applied in industry due to low yields.

Porphyrin-iron complexes (hemes) are well known tightly-bound cofactors (prosthetic groups) in a number of metalloproteins – Hemoglobin, Cytochrome C, P450, Peroxidases, etc. The biological functions that they are involved in include dioxygen transportation, electron transfer, and redox chemistry. The heme enzymes and synthetic complexes require an additional cofactor in order to activate the $O_2$ molecule for substrate oxidation. Two Porphyrin-iron complexes however can bind dioxygen as a peroxide in a Di-iron complex and cleave the O-O bond, as reported by Chin and...
co-workers [93]. The O-O bond cleavage in binuclear Porphyrin peroxy complexes was recently studied theoretically [94] with DFT. The stability of the binuclear center at normal temperature was explained with the splittings of the spin states and the way they are changed in the presence of axial ligands.

In the present work, the stepwise mechanism of the catalytic oxidation of cyclohexane to adipic acid was investigated theoretically. Hybrid density functional theory was used to model the putative oxidative stages in the adipic acid synthesis. The reaction was modelled according to the conditions reported in ref. [92] – solvent-free reaction system with molecular oxygen as the only oxidant, and a heme Fe-oxo complex (PorfFe(IV)=O) as a catalyst, at 413K temperature.

11.2 Computational technique.

The common computational technique as described in Chapter 5 of Part 2 was applied in this study. The modified B3LYP* functional was used for checking the reliability of the results. The reduction of the Hartree-Fock exchange to 15% resulted in the typical stabilization of the low-spin states, and destabilization of the high-spin states. The relative energies reported below are the ones obtained with the original B3LYP.

The solvent was modeled as a macroscopic continuum with a dielectric constant $\varepsilon=2.0$ (corresponding to cyclohexane). Since the model used in the study has zero total charge and basically no charge separation happens, the solvent effects are very small, in the range of 1–2 kcal/mol. Reported spin populations are used to indicate the spin- and oxidation state of the metal ion, and were derived from a Mulliken population analysis.

The complex used as a catalyst in the experiment is meso-Tetra(o-chlorophenyl)iron Porphyrin (T(o-Cl)PPFe). Unsubstituted Porphyrin ligand was used in the calculations. This strategy for decreasing the computational costs has been proven to give adequate results in recent theoretical study where the energetic differences between substituted and unsubstituted Porphyrin models were found insignificant [94].

Since for loosely bound intermediates the harmonic approximation gives poor results, the thermodynamic corrections for such structures were taken from the infinitely separated species. Although the partition functions for an ideal gas are very accurate for predicting gas phase entropies, errors are introduced when this method is applied to solvent molecules. In particular, the calculated gas phase entropy of the cyclohexane molecule is higher than the entropy it has in the liquid phase. In principle, it is possible to estimate the difference by using experimentally determined entropies, and for cyclohex-
11.2. Computational technique.

It corresponds to the measured vaporization enthalpy of 9.3 kcal/mol at T=413K [95, 96]. This number is very close to the calculated entropy lost of 8.3 kcal/mol (at T=413K) for the transition state for hydrogen atom transfer from cyclohexane to Fe(IV)-oxo. In the later reaction steps molecules very similar to the cyclohexane form complexes and react with the heme complex, but they should be present in low concentrations in the solution and the entropy loss associated with complex formation should thus be larger than for cyclohexane. At present, we are not certain about the accuracy of the calculated entropy affects, therefore the relative enthalpies was used in the energy profiles.

To account for the Van der Waals forces additional calculations were done with B3LYP-D using the Orca Package [83, 84]. Dispersion effects of 2 to 4 kcal/mol were calculated for the formation of a complex between the substrate molecules and the catalyst. However, the effect on the relative energies between different H-bonded complexes and transition states were found to be smaller, up to 1.0 kcal/mol.

![Scheme 11.1](scheme.png)

Scheme 11.1: Schematic representation of the electronic mechanism for the hydrogen atom abstraction, leading to AFM and FM triplet and quintet states.
11.3 Suggested reaction mechanism

The oxidation of cyclohexane to adipic acid by the biomimetic heme complex T(o-Cl)PP-Fe(IV)-O was studied using DFT. Experimentally the reaction was done in cyclohexane as a solvent, and with molecular oxygen as the only oxidant. The suggested mechanism consists of three major steps. In the first one the cyclohexane is hydroxylated to 1,2-cyclohexanediol. It should be noted here that the cyclohexane molecule is in its lowest-energy chair conformation, where axial and equatorial positions can be distinguished. The two hydroxylations of the cyclohexane, described below, occur at neighboring (ortho) equatorial position. In the second major step 1,2-cyclohexanediol is oxidized to 1,6-hexanedial with opening of the ring, and in the final step both aldehyde groups of 1,6-hexanediol are oxidized, and the adipic acid is formed. These three main reactions share some common features. The ground spin state of the reactants and the products is triplet (S=2). This is determined by the spin splitting for the Fe(IV)-oxo and Fe(II) heme complexes, which both have a triplet ground state. For every oxidation however, the reactive state was found to be the quintet (S=4) – the system undergoes a spin crossing from triplet to quintet state before the first transition state, and when the product is formed another spin crossing back to the triplet occurs. Every oxidative step starts with hydrogen atom abstraction, resulting in the formation of Fe(III)-OH and a radical. Depending on the spin coupling between Fe(III)-OH and the radical molecule, two different triplet and quintet states exist – Antiferromagnetically (AFM) and Ferromagnetically (FM) coupled (Scheme 11.1). Details about all the different spin states that were calculated during the investigation can be found in Paper V.

11.3.1 Oxidation of cyclohexane to 1,2-cyclohexanediol

The oxidative hydroxylation starts with Hydrogen Atom Transfer (HAT) from cyclohexane to the Fe(IV)-oxo species, leading to a radical intermediate [(Porph)Fe(III)(OH)(C₆H₁₁)•]. In the next step a rebound of the OH group occurs to form the alcohol. The ground state of the reactant complex [(Porph)Fe(IV)(O)(C₆H₁₂)] is a triplet (Figure 11.1 A) having formally two unpaired electrons on iron, one of which is partially delocalized to the oxo group (S_{Fe}=1.21 and S_{oxo}=0.83, where S_X refers to the Mulliken spin population of fragment X). The lowest barrier for HAT, however, is found in the quintet state. In the reactant complex, the quintet state is 7.8 kcal/mol higher than the triplet ground state. But typically for the Fe(IV)-oxo group, the system undergoes a spin crossing before the transition state is reached. In the TS for HAT the quintet state is 3.6 kcal/mol lower than the AFM triplet state. The enthalpy barrier determined by the TS in the quintet state (Figure 11.1 B) is 13.6 kcal/mol with respect to the starting complex.
11.3. Suggested reaction mechanism

\[(\text{Porph})\text{Fe(IV)}(\text{O})(\text{C}_6\text{H}_{12})\] in its ground triplet state. The HAT is endergonic with 6.7 kcal/mol.

In the intermediate complex Fe has formally been reduced from Fe(IV) to Fe(III). The ground state of the intermediate is a quintet, in which iron has five unpaired electrons \((S_{Fe}=4.00, S_{OH}=0.39)\) AFM coupled to the unpaired electron on the substrate radical. The AFM coupled triplet state is 5.9 kcal/mol higher in energy. The enthalpy of formation of the closed-shell product cyclohexanol is -42.0 kcal/mol and the barrier for this process is expected to be very low. The optimized quintet TS is 0.7 kcal/mol above the preceding intermediate complex \([(\text{Porph})\text{Fe(III)}(\text{OH})(\text{C}_6\text{H}_{11})^*]\). In the reactive quintet state the second step is exergonic by 37.0 kcal/mol. However, before the product complex \([(\text{Porph})\text{Fe(II)}(\text{C}_6\text{H}_{11}(\text{OH}))]\) is reached, a second spin crossing occurs and the ground state of the product complex is the triplet state. The product complex is 0.7 kcal/mol more stable than the free species.

Concerning the open-shell singlet state, the singlet radical intermediate \([(\text{Porph})\text{Fe(III)}(\text{OH})(\text{C}_6\text{H}_{11})^*]\) was found to lie 9.9 kcal/mol above the ground quintet state. Therefore it was concluded that the singlet state is not involved in the reaction and transition states were not located for this spin state.

The second hydroxylation starts with a new \([(\text{Porph})\text{Fe(IV)}(\text{O})]\) reactant, which forms a weak hydrogen-bonded complex (binding enthalpy -8.3 kcal/mol) with the OH-group of cyclohexanol \(\text{C}_6\text{H}_{11}(\text{OH})\) (Figure 11.2 A).
The spin splitting for this reactant shows a triplet ground state ($S_{Fe}=1.28$ and $S_{oxo}=0.78$), quintet state lying 7.7 kcal/mol higher, and singlet at 9.5 kcal/mol (Figure 11.3).

The hydrogen atom that is abstracted in this stage is the one at ortho-equatorial position, relative to the OH-group of the cyclohexanol (see the lowest optimized transition state in Figure 11.2 B). The same electronic aspects of the HAT mechanism can be considered, as in the first oxidation. Two possible radical intermediates [(Porph)Fe(III)(OH)(C₆H₁₀OH)•], AFM- and FM-coupled, arise from both the triplet and the quintet reactants. The enthalpy barrier for formation of the AFM triplet is 20.8 kcal/mol, while for the AFM quintet it is 3.0 kcal/mol lower, and is the lowest found. It is seen that again a spin crossing to the quintet state occurs before the HAT transition state. The HAT process is endergonic with 8.6 kcal/mol and the ground state of the resulting intermediate [(Porph)Fe(III)(OH)(C₆H₁₀OH)•] is an AFM quintet ($S_{Fe}=4.03$, $S_{OH}=0.32$), while the the FM triplet is 3.2 kcal/mol higher. The hydroxylation of the radical proceeds on the AFM quintet surface, but the ground state for the product [(Porph)Fe(II)(C₆H₁₀(OH)₂)] is again the triplet. The lowest optimized TS (AFM quintet) is 3.1 kcal/mol above the preceding intermediate. The free energy for the formation of the product complex is -35.4 kcal/mol, determined by the energy of the ground triplet state. This complex is 0.9 kcal/mol more stable compared to the separate products. The resulting 1,2-
11.3. Suggested reaction mechanism

11.3.2 1,2-cyclohexanediol oxidation and ring cleavage

The oxo group of [(Porph)Fe(IV)(O)] forms a weakly (ΔH=-4.0 kcal/mol) hydrogen bonded complex [(Porph)Fe(IV)(O)(C₆H₁₀(OH)₂)] with the two OH groups of 1,2-cyclohexanediol C₆H₁₀(OH)₂. The ground state of the reactant complex (Figure 11.3 A) is a triplet $S_{Fe}=1.42$ and $S_{oxo}=0.65$, with the quintet and the singlet at 5.8 and 9.9 kcal/mol, respectively (Figure 11.4).

First a hydrogen atom is transferred from one of the OH groups of the substrate to the oxo group of [(Porph)Fe(IV)(O)]. This yields a H-bonded intermediate complex [(Porph)Fe(III)(OH)(C₆H₁₀(OH)O*)] (Figure 11.3 C), whose ground state is a quintet with five unpaired electrons in iron, antiferromagnetically coupled to the unpaired electron on the substrate radical. The AFM triplet and the FM quintet states are quite close, at 4.6 and 5.0 kcal/mol above the ground state, respectively. The spin splitting for the HAT transition

Figure 11.2: Optimized structures of A: the reactant complex [(Porph)Fe(IV)(O)(C₆H₁₁OH)] in its ground triplet state; B: HAT transition state in AFM quintet state. The most important distances and spin populations are given.
11.3. Suggested reaction mechanism

Scheme 11.3: Energy diagram for the second hydroxylation.

The intermediate complex, in which the substrate is in a radical state, is highly unstable with respect to formation of the closed-shell product 1,6-hexanediol. The free energy for this step is -39.7 kcal/mol and consequently the barrier is very low. The electronic energy barrier for this step in the gas phase is only 3.5 kcal/mol. When the dielectric solvent effects and the enthalpy correction are added, the energy of the optimized TS becomes 5.4 kcal/mol lower than the preceding intermediate complex [(Porph)Fe(III)(OH)(C₆H₁₀(OH)O•)]. The transition state for C-C bond dissociation is concerted with the HAT from the second OH group on the substrate radical to the OH group on iron, see Figure 11.3 D. A sequential process was also considered. It is not possible that the C-C bond dissociates before the second HAT since this requires that the substrate enters a di-radical state. Both the triplet and open-shell singlet states of this di-radical are very high in energy (ca. 40 kcal/mol higher than the doublet state of the mono radical in the intermediate complex). In all attempts to find a TS for the second HAT, the C-C bond dissociates with the OH bond. It can be concluded that the first HAT is close to concerted with the C-C bond dissociation, and the second HAT (Scheme 11.4). In the reactive quintet state the second step is exergonic by 37.9 kcal/mol. However, before the H-bonded product complex
Scheme 11.4: Energy diagram for the two HAT and ring opening of the 1,2-cyclohexanediol ring.

\[
[(\text{Porph})\text{Fe(II)}(\text{OH}_2)(\text{OHCC}_4\text{H}_8\text{CHO})]\text{ is reached, a second spin crossing occurs and the ground state of the product complex is the triplet state. The H-bonded product complex is in close equilibrium with the free species } (\Delta H=-3.5 \text{ kcal/mol}).
\]

During the whole reaction, the closed-shell singlet state is the highest state calculated. The closed-shell singlet state is thus not involved in the reaction.

11.3.3 Oxidation of adipic aldehyde to adipic acid

The last major stage consists of two consecutive oxidations of the carbonyl groups of the adipic aldehyde to carboxylic groups. Essentially, each oxidations is again a hydrogen atom transfer followed by a rebound of the OH group. The chemical steps have the same features as in the first two major steps: triplet ground state for the reactant and product; spin crossing before the HAT transition state and after the OH rebound; AFM quintet reactive state.

The reactant is a loose complex between the oxidant \([(\text{Porph})\text{Fe(IV)}(\text{O})]\) and the acyclic adipic aldehyde with binding enthalpy -0.7 kcal/mol (Figure 11.4 A). In the first step a carbonyl hydrogen atom is transfered from one of the carbonyl groups of the substrate to the oxo group of the metal complex. On the reactive AFM quintet PES the calculated activation barrier is 7.7
11.3. Suggested reaction mechanism

kcal/mol with respect to the triplet reactant, and the reaction is exergonic with 7.1 kcal/mol (optimized transition state in Figure 11.4 B). The following OH rebound goes through an activation barrier of 3.2 kcal/mol and the exergonicity (determined from the triplet ground state of the product) is 50.2 kcal/mol (Scheme 11.5).

In the very last HAT the activation barriers for the AFM quintet and AFM triplet spin states for the HAT transition state differ by only 0.8 kcal/mol (Scheme 11.6). Still the quintet should be considered as a reactive state with activation barrier of 8.2 kcal/mol. The reactant and the HAT transition state structures are shown in Figure 11.5. The resulting radical intermediate lies 3.6 kcal/mol below the ground state reactant, and it takes 1.8 kcal/mol for the OH rebound to occur. The adipic acid is formed with exergonicity of 51.2 kcal/mol (triplet ground state).

The catalytic formation of adipic acid from cyclohexane and dioxygen was found to be a strongly exergonic process. Each of the sequential steps, but one, consists of hydrogen abstraction and subsequent hydroxide rebound. The exception is the ring cleavage step where two sub-sequential hydrogen abstractions occur. In all of the major steps hydrogen abstraction is rate limiting. However the overall rate limiting HAT for the whole reaction is the one for the hydroxylation of cyclohexanol to ortho-cyclohexanediol. The calculated enthalpy of activation for this step is 17.8 kcal/mol. The entropy effect is not expected to increase this number significantly. The reason that this particular HAT becomes rate limiting can be found in the formation of a rather stable [(Porph)Fe(II)(C_6H_{10}(OH)_2)] adduct prior to the HAT. This step is possibly crucial in differentiating between competing reactions, that might lead to the formation of meta- or para-cyclohexanediol, which in turn would lead to different final products. These alternative paths have not been considered in this work, and the challenge to explore them remains. The presented mechanism shows, however, that the underlying chemistry in itself is rather basic.
11.3. Suggested reaction mechanism

Figure 11.3: Optimized structures of A: the reactant complex \([\text{Porph}\text{Fe(IV)}(\text{O})(\text{C}_6\text{H}_{10}\text{(OH)}_2)]\) in its ground triplet state; B: HAT transition state in AFM quintet state; C: AFM quintet state for the intermediate complex \([(\text{Porph})\text{Fe(III)}(\text{OH})(\text{C}_6\text{H}_{10}\text{(OH)}\text{O}^*)]\); D: transition state for the second HAT and C-C bond cleavage. The most important distances and spin populations are given.
Scheme 11.5: Energy diagram for the oxidation of the first carbonyl group of the adipic aldehyde.

Figure 11.4: Optimized structures of A: the reactant complex [(Porph)Fe(IV)(O)(OHCC₄H₈CHO)] in its ground triplet state; B: HAT transition state in AFM quintet state. The most important distances and spin populations are given.
Scheme 11.6: Energy diagram for the final oxidative step leading to adipic acid.

Figure 11.5: Optimized structures of A: the reactant complex [(Porph)Fe(IV)(O)(OHCC₄H₈COOH)] in its ground triplet state; B: HAT transition state in AFM quintet state. The most important distances and spin populations are given.
Part V:
Concluding Remarks

The investigations presented in this thesis aim at bringing some insights and better understanding of the catalytic mechanism of ortho-hydroxyl dioxygenases, and dioxygen activation in general. A hybrid DFT method is used to model three enzymes (Homoprotocatechuate 2,3-dioxygenase from Brevibacterium fuscum Bf 2,3-HPCD, Manganese-Dependent Homoprotocatechuate 2,3-Dioxygenase MndD, and Homogentisate Dioxygenase HGD), and two biomimetic complexes. The reported results are obtained with models consisting of up to 134 atoms.

The experimentally observed similarities between Bf 2,3-HPCD and MndD are reproduced in the calculations, and the activation barriers obtained are in good agreement with the experiment. The suggested mechanism for the catalytic reaction of HGD agrees with several experimental observations for extradiol dioxygenases. The alternative reaction paths were found unprobable by reason of higher barriers.

The calculations for the non-heme iron complex show that only the quartet state is likely to be involved in an energetically plausible mechanism. Even though structurally different from the typical intradiol enzyme active site, the complex uses a mechanism with the same major features – formation of bridging peroxide and homolytic O-O bond cleavage. The formation of the bridging peroxide was found rate limiting in the proposed mechanism.

Based on the computational results, a reaction mechanism was suggested for the conversion of cyclohexane to adipic acid in the presence of molecular oxygen, and a biomimetic heme complex as a catalyst. The mechanism consists of several consecutive oxidative steps, where a hydrogen atom transfer process was found rate limiting. The overall rate limiting HAT is possibly important for differentiation between the desired and the side products.
Part VI: Acknowledgements

First of all, I would like to thank my supervisors Margareta and Per for accepting me in the group and giving me the opportunity to work in this area. I have found friendly atmosphere and a lot of help. I am very grateful to Tomek, for guiding my first steps, and being such a good friend.

Many thanks to all my former and present colleagues for their help and friendship. Special thanks to Holger (from Germany) and Johannes, for the great times we had and for our collaboration. Thanks to Martin, Rossen and Holger for reading the thesis.

I am grateful to all my previous supervisors from the University of Sofia, Bulgaria. They introduced me to Theoretical Chemistry and helped me in difficult moments.

Warm thanks to my family for the fact that I exist, and for their support all the time, in all the difficult for us moments. Thanks to Polina’s family for their hospitality and care.

Finally I would like to thank Polina, Monica and Filip – my present family. Thank you Polina for your love and support, for making me happy, and thank you Monica and Filip, for bringing me all the joy! I love you!

THANK YOU ALL!


Att studera olika enzymers struktur och deras reaktionsvägar är ett sätt att försöka förstå deras stora katalytiska kapacitet. Sådan kunskap är av stor betydelse inom alla områden där katalytiska egenskaper utnyttjas, som till exempel i läkemedelsindustrin och kemiindustrin.


[41] T. Vreven; K. N. Kudin; J. C. Burant; J. M. Millam; S. S. Iyengar; J. Tomasi; V. Barone; B. Mennucci; M. Cossi; G. Scalmani; N. Rega; G. A. Petersson; H. Nakatsuji; M. Hada M. Ehara; K. Toyota; R. Fukuda; J. Hasegawa; M. Ishida; T. Nakajima; Y. Honda; O. Kitao; H. Nakai; M. Klene; X. Li; J. E. Knox; H. P. Hratchian; J. B. Cross; V. Bakken; C. Adamo; J. Jaramillo; R. Gomperts; R. E. Stratmann; O. Yazyev; A. J. Austin; R. Cammi; C. Pomelli; J. W. Ochterski; P. Y. Ayala; K. Morokuma; G. A. Voth; P. Salvador; J. J. Dannenberg; V. G. Zakrzewski; S. Dapprich; A. D. Daniels; M. C. Strain; O. Farkas; D. K. Malick; A. D. Rabuck; K. Raghavachari; J. B. Foresman; J. V. Ortiz; Q. Cui; A. G. Baboul; S. Clifford; J. Cioslowski; B. B. Stefanov; G. Liu; A. Liashenko; P. Piskorz; I. Komaromi; R. L. Martin; D. J. Fox; T. Keith; M. A. Al-Laham; C. Y. Peng; A. Nanayakkara; M. Challacombe; P. M. W. Gill; B. Johnson; W. Chen; M. W. Wong; C. Gonzalez; M. J. Frisch; G. W. Trucks; H. B. Schlegel; G. E. Scuseria; M. A. Robb; J. R. Cheeseman; J. A. Montgomery; Jr. and Wallingford CT J. A. Pople; Gaussian Inc. Gaussian 03, revision d01. 2007.


